ARTICLE

ENGAGING FACTS AND POLICY: A MULTI- INSTITUTIONAL APPROACH TO PATENT SYSTEM REFORM

Arti K. Rai*

The Court of Appeals for the Federal Circuit, charged with adjudicating appeals in patent cases, has adopted an unusual approach that arrogates power over fact finding while it simultaneously invokes rule-formalism. Although the Federal Circuit's approach may be justified by the fact-finding and policy application deficiencies of the trial courts and the Patent and Trademark Office (PTO), it has had a negative impact on innovation policy and has resulted in a patent system that is sorely in need of reform. This Article argues that because of the interdependence of the various institutions within the patent system, reform of the system must be both multi-institutional and closely attentive to the institutional competence of the system's actors. Although Congress should clearly bolster the PTO's fact-finding abilities, giving plenary responsibility over factual questions to the PTO would not be cost effective. Accordingly, Congress should endow the system with improved fact-finding expertise through the institution of specialized trial courts. As for actual policy formulation, each of the available institutional options—the legislature, the PTO, and the courts—has substantial associated liabilities. On balance, however, the Federal Circuit is probably best positioned to formulate patent policy, so long as the fact-finding expertise of inferior institutions is bolstered and additional appellate mechanisms are instituted. By paying attention to institutional design and revising our institutions accordingly, we can achieve the patent system we should have had all along.

* Assistant Professor, University of Pennsylvania Law School. I thank Michael Abramowicz, Stuart Benjamin, Dan Burk, Michael Carrier, Paul Carrington, Eric Claeys, Rochelle Dreyfuss, John Duffy, Rebecca Eisenberg, Susan Freiwald, Larry Helfer, Bob Hunt, Orin Kerr, Seth Kreimer, Mark Lemley, Cecil Quillen, Ed Rubin, Polk Wagner, Phil Weiser, and Galen Sherwin for helpful comments. Earlier versions of this paper were presented at the Spring 2002 Federal Trade Commission/Department of Justice Hearings on the Patent System, the University of Michigan Law School, Duke University Law School, the University of Minnesota Law School, and the University of Pennsylvania Law School. I thank the participants at those symposia for helpful comments.

TABLE OF CONTENTS

INTRODUCTION	1036
I. THE FEDERAL CIRCUIT AS FACT FINDER	1042
A. The Law/Fact Distinction	1042
B. The Role of Facts in Patent Law.....	1044
1. The Role of Facts in Patent Scope.....	1046
2. The Role of Facts in Patent Validity: Nonobviousness and Disclosure.....	1049
C. Federal Circuit Review of PTO Fact Finding	1052
1. Direct Review of PTO Patent Denials	1052
2. Review of PTO Patent Grants	1056
D. Federal Circuit Review of Trial Court Fact Finding ...	1057
1. Review of Fact Finding in Trial Court Claim Construction.....	1057
2. Review of Fact Finding in Trial Court Nonobviousness and Disclosure Decisions.....	1060
II. THE NORMATIVE ANALYSIS: SHOULD THE FEDERAL CIRCUIT EXERCISE DE NOVO REVIEW?	1065
A. Judicial Deference to PTO Fact Finding	1066
1. General Considerations	1066
2. Should the Federal Circuit Defer?	1068
B. Deference to Trial Courts	1086
1. General Considerations	1086
2. Should the Federal Circuit Defer?	1088
C. The Relationship Between Fact Finding and Policy Application	1101
III. FILLING THE POLICY GAP	1102
A. The Federal Circuit: A Study in Formalism	1103
B. Formalism or Capture?	1110
C. The Limits of Formalism in Patent Law	1115
1. Formalism Generally: Pro and Con	1115
2. Structural Arguments Against Patent Formalism ..	1116
3. Functional Arguments Against Patent Formalism..	1120
D. Institutional Structure and Resources: The Comparative Analysis.....	1122
1. The Comparison to Legislative Policymaking	1127
CONCLUSION	1134

INTRODUCTION

As the commodity of information has become increasingly vital to the national and global economy, intellectual property has elbowed its way from the legal periphery to center stage. Nowhere can we see the expanding reach of intellectual property more clearly than in the growth of the patent system. In recent years, the number of patents issued annu-

ally has increased precipitously.¹ The growing prominence of the patent system has revealed, however, an institutional structure in need of repair. The Patent and Trademark Office (PTO), the administrative agency responsible for issuing patents, is not designed to evaluate properly the technology in many of the patent applications that it receives. Federal district courts, which are the initial arbiters of private party disputes over patent scope, validity, and infringement, have flaws that are equally fundamental. Trial court decisions are dependent not only on judges who may be unversed in the relevant technology but also upon “black box” determinations by lay juries. For its part, the Court of Appeals for the Federal Circuit (CAFC or Federal Circuit), which Congress set up in 1982 to hear all appeals in patent cases,² has pursued two major strategies that appear anomalous: first, relatively vigorous de novo review, including de novo review of questions of technological fact; and second, the adoption of formalist, bright-line rules that leave inferior decisionmakers little room for factual inquiry.³ De novo review of factual questions by an appellate court is obviously anomalous. The court’s adoption of bright-line rules that are insensitive both to technological fact and to related issues of innovation policy is also suspect: As I will argue below, Congress has chosen to make the requirements of patentability fact-specific and the

1. In the calendar year 2001, the Patent and Trade Office (PTO) issued a total of 166,039 utility patents. This is approximately a sixty-three percent increase over calendar year 1995. See Information Products Division/Technology and Forecast Branch, U.S. Patent and Trademark Office, U.S. Patent Statistics, Calendar Years 1963–2001 (Mar. 2002), at http://www.uspto.gov/web/offices/ac/ido/oeip/taf/us_stat.pdf (on file with the *Columbia Law Review*). In addition, the number of utility patent applications received by the PTO in the year 2001 was 326,508. *Id.*

2. The Federal Circuit hears patent appeals in two contexts. The first context is direct appeal from a PTO denial of patentability. The second context is appeal from a regional trial court decision. In the latter context, a patent has been granted, and the patent owner has brought an action against an alleged infringer (or a possible infringer has brought a declaratory judgment action challenging patent validity). The Federal Circuit also hears appeals of other types of cases, including the following: PTO trademark decisions; final decisions of the Court of International Trade; certain types of technology transfer litigation; non-tax and non-tort government contract cases; and decisions from the Merit Systems Protection Board. Rochelle Dreyfuss, *The Federal Circuit: A Case Study in Specialized Courts*, 64 N.Y.U. L. Rev. 1, 4 & nn.22–26 (1989) [hereinafter Dreyfuss, *Federal Circuit*]. This jurisdictional range notwithstanding, the reason for setting up the Federal Circuit was the view that the regional courts of appeal were creating a body of patent law that was inconsistent and unpredictable. According to proponents of a specialized court, this inconsistency and unpredictability undermined the value of patents and led to forum shopping. See generally Comm’n on Revision of the Federal Court Appellate System, *Structure and Internal Procedures: Recommendations for Change*, reprinted in 67 F.R.D. 195, 217–21 (1975) (suggesting that forum shopping under the old system demeaned “the entire judicial process,” and the patent system in particular).

3. Where possible, the Federal Circuit invokes both strategies: It overturns the inferior decisionmaker’s fact finding and replaces it with a bright-line rule that will obviate the need for future inquiry into fact. See *infra* text accompanying notes 74–79.

language of the statute open to policy-oriented interpretation by the judiciary.⁴

Legal scholars and commentators have identified and discussed some of the difficulties faced by the individual institutions within the patent system.⁵ But they have not incorporated into their analyses a recognition of all of the relevant institutions and their highly interdependent roles. Most notably, though a few scholars have alluded to the Federal Circuit's tendencies towards *de novo* fact finding and bright-line rules,

4. Notably, because the Supreme Court, at least historically, has rarely reviewed Federal Circuit cases, the Federal Circuit's formalist jurisprudence has typically constituted the final word on patent questions. See generally Mark D. Janis, *Patent Law in the Age of the Invisible Supreme Court*, 2001 U. Ill. L. Rev. 387 (noting that Federal Circuit has effectively been the Supreme Court of patents). In the last few years, by contrast, the U.S. Supreme Court has shown significant interest in such review. See *infra* text accompanying notes 8–9, 393–395.

5. Much of the literature has analyzed deficiencies in the PTO process for evaluating patent applications. E.g., Mark D. Janis, *Rethinking Reexamination: Toward a Viable Administrative Revocation System for U.S. Patent Law*, 11 Harv. J.L. & Tech. 1, 3 (1997); Jay P. Kesan & Marc Banik, *Patents as Incomplete Contracts: Aligning Incentives for R&D Investment with Incentives to Disclose Prior Art*, 2 Wash. U. J.L. & Pol'y 23, 24 (2000); Robert P. Merges, *As Many as Six Impossible Patents Before Breakfast: Property Rights for Business Concepts and Patent System Reform*, 14 Berkeley Tech. L.J. 577, 589–615 (1999) [hereinafter Merges, *Six Impossible Patents*]; Cecil D. Quillen, Jr. & Ogden H. Webster, *Continuing Patent Applications and Performance of the U.S. Patent and Trademark Office*, 11 Fed. Cir. B.J. 1, 3 (2001) (citing ninety-five percent ultimate approval rate for patent applications in U.S. PTO, if continuing applications are counted, as contrasted with sixty-eight and seventy-five percent allowance rates for European and Japanese Patent Offices respectively); John R. Thomas, *Collusion and Collective Action in the Patent System: A Proposal for Patent Bounties*, 2001 U. Ill. L. Rev. 305, 305. Scholars have also discussed the limitations of generalist trial court judges and juries in deciding patent cases. E.g., John B. Pegram, *Should There Be a U.S. Trial Court with a Specialization in Patent Litigation?*, 82 J. Pat. & Trademark Off. Soc'y 765, 790 (2000). A few articles have considered the refusal of the Federal Circuit to show deference to the PTO's factual determinations. E.g., Craig Allen Nard, *Deference, Defiance, and the Useful Arts*, 56 Ohio St. L.J. 1415, 1421 (1995); Arti Rai, *Addressing the Patent Gold Rush: The Role of Deference to PTO Patent Denials*, 2 Wash. U. J.L. & Pol'y 199, 201 (2000) [hereinafter Rai, *Patent Gold Rush*]. One recent article has considered the Federal Circuit's "hyperactivity" in deciding *de novo* inquiries like infringement that the court itself has recognized to be factual in nature. See William C. Rooklidge & Matthew F. Weil, *Judicial Hyperactivity: The Federal Circuit's Discomfort with its Appellate Role*, 15 Berkeley Tech. L.J. 725 *passim* (2000). Unlike Weil and Rooklidge, I also examine cases where the Federal Circuit has recharacterized questions of fact as questions of law. Cf. *id.* at 750 (noting that "[j]udicial expansion involving . . . the labeling of issues is . . . beyond the scope of this article"). Because legal principles have precedential value in future cases, such recharacterization of fact as law is ultimately more important than the *de novo* review to which Weil and Rooklidge point. Further, many of the cases that Rooklidge and Weil discuss, in which the Federal Circuit has decided factual questions such as infringement *de novo*, are cases in which the court first characterized as a question of law (and hence reviewed *de novo*) the district court's claim construction. After deciding on a different claim construction than the district court, the Federal Circuit then used that construction to declare that there could be no factual dispute on the question of infringement. See *infra* notes 108–109 and accompanying text.

they have not discussed whether this behavior may be justified by the court's dependence on inferior decisionmakers of questionable competence in the realms of fact finding and factually oriented policy application.⁶ Because of this interdependence, patent reform requires multi-institutional analysis. This multi-institutional analysis must also be comparative in nature: Only by evaluating the relative competence of the various institutions in performing the tasks required by the patent process can we hope to design a system that works reasonably well—or, at a minimum, less imperfectly than the alternatives.⁷

The need for multi-institutional analysis, with a particular focus on the proper role of the Federal Circuit, is highlighted by the recent interest of the Supreme Court in reviewing cases from the Federal Circuit. Some of these cases have involved disputes over allocation of power between the Federal Circuit and other institutions, both within and outside the patent system.⁸ Other cases appear to reflect a concern that the Federal Circuit is paying insufficient attention to questions of innovation policy.⁹

In conducting this multi-institutional analysis, I do not presuppose a particular economic theory of patents. Even those economic theorists who believe in an expansive role for patents (typically on the assumption that patent rights are somewhat similar to rights in ordinary property and thereby serve primarily as incentives for efficient development and commercialization rather than as incentives to invent¹⁰) do not argue that we should allow patent rights on products or processes that are already in the public domain. Many of the patents issued by the PTO appear to cover inventions that are effectively in the public domain.¹¹ More generally, because this Article's focus is institutional, I strive to steer clear of assumptions regarding whether achieving the patent system's constitutionally authorized innovation goals¹² requires a "low" or "high" level of

6. Throughout this Article, I focus on possible justifications for the Federal Circuit's anomalous approach to questions of fact and policy. For my purposes, justification, not motivation, is the relevant issue. To the extent that any justification for the Federal Circuit's anomalous approach is removed, it can and should be forced to change its approach.

7. See generally Neil K. Komesar, *Imperfect Alternatives: Choosing Institutions in Law, Economics, and Public Policy* 4 (1994) (arguing that "comparative institutional analysis" is essential to understanding law and public policy generally).

8. See, e.g., *Holmes Group, Inc. v. Vornado Air Circulation Sys., Inc.*, 535 U.S. 826 (2002) (limiting "arising under" jurisdiction of Federal Circuit).

9. See, e.g., *Festo Corp. v. Shoketsu Kinzoku Kogyo Kabushiki Co.*, 535 U.S. 722 (2002) (reversing Federal Circuit's interpretation of prosecution history estoppel). By innovation policy, I mean policy directed towards promoting innovation.

10. See, e.g., Edmund W. Kitch, *The Nature and Function of the Patent System*, 20 *J.L. & Econ.* 265, 265 (1977) (arguing that the patent system performs the function of "increas[ing] the output from resources used for technological innovation").

11. See, e.g., Merges, *Six Impossible Patents*, *supra* note 5.

12. See U.S. Const. art. I, § 8, cl. 8 (authorizing patent grants to promote the "Progress of Science and useful Arts").

patent protection. Rather, in discussing which institution is best placed to address legal and policy issues, I focus on the extent to which the institution is likely to hear from, and be receptive to, a wide variety of arguments on how innovation is best promoted. To put the point another way, although I assume a normative goal of innovation,¹³ I argue that the test of whether an institution is well placed to promote that goal is its openness to arguments by innovation advocates on all sides.

The ostensibly discrete problems of our patent institutions stem from an initial mistake in institutional design. This mistake was made in 1982, when Congress implemented patent reform by focusing on the appellate level and attempting to concentrate expertise there. The central problems of our patent system follow inexorably from Congress's 1982 decision. The Federal Circuit's exercise of *de novo* review over fact finding may be rationalized by the argument that, despite the court's appellate status, it still has fact-finding capabilities superior to those of the other decisionmakers in the patent system. Indeed, given Congress's decision to focus on the appellate level, this argument has a substantial foundation. Generalist trial judges, and the juries empanelled by trial judges, may be overwhelmed by the technology involved in patent cases.¹⁴ The PTO also has well-documented limits in its procedures for evaluating technology. Nonetheless, appellate courts are hardly equipped for fact finding. Thus, the consequence of this institutional design is, at a minimum, inefficiency and technologically questionable decisions in individual cases. More broadly, the Federal Circuit has substituted formalist decisionmaking for the fact-specific, policy-oriented analysis that is required by the open-ended language of the patent statute: Once again, although the court's approach might be justified by reference to the problems of the inferior decisionmakers, it is far from optimal. Bright-line rules are, in reality, not always easy to implement. In addition, bright-line rules cannot be expected to do a good job of furthering the innovation goals of patent law. In any event, even if certain bright-line rules could incorporate innovation policy goals to some extent, the rules that the Federal Circuit has chosen do not appear to do so.

In this Article, I argue that multi-institutional reform must turn on a recognition of two institutional deficiencies: First, no institution within the patent system has the expertise necessary for finding facts accurately; and second, no institution has taken responsibility for elaborating patent law in the fact-specific, policy-oriented manner that the language of the

13. Cf. Einer R. Elhauge, *Does Interest Group Theory Justify More Intrusive Judicial Review?*, 101 *Yale L.J.* 31, 48-59 (1991) (arguing that institutional analysts should specify a normative baseline).

14. In a prescient article written in 1989, soon after the establishment of the Federal Circuit, Rochelle Dreyfuss alluded briefly to this potential problem with the establishment of a specialized court at the appellate level. She observed that if a specialized *trial* court had been set up, "the superior fact finding capability of the special court would have been fully exploited and the assumptions underlying the normal standards of review maintained." Dreyfuss, *Federal Circuit*, *supra* note 2, at 73-74.

statute encourages. The latter deficiency is particularly notable: Although commentators have not focused on the issue, the history of the patent statute as well as its language strongly suggest that Congress has delegated policymaking responsibility in patent law to the judiciary.

How these institutional deficiencies should be addressed presents an interesting but difficult question. Under traditional principles of administrative law, the agency is endowed with the expertise necessary for fact finding. Courts then defer to this fact finding. As I argue below, these principles have some applicability in the patent arena. Congress should indeed advance the PTO beyond its nineteenth-century roots and make it an agency capable of sorting out the complex factual details of various important technologies. But it would be unduly expensive for the PTO to deploy the full range of expertise in the context of every patent grant. Thus it is a good idea to have a second layer of fact-finding expertise at the level of the trial court. In addition, with greater fact-finding expertise, the PTO and the trial courts would both presumably be better able to apply policy directives: After all, application of such policy directives often requires inquiry into case-specific factual scenarios.¹⁵ If greater fact-finding and policy application expertise were vested at the administrative and trial court levels, the role of appellate review within the patent system could substantially be reconceived. To a much greater extent than is currently the case, appellate review could be informed by generalist input. In addition, such review could focus on developing the open-ended language of the patent statute in the larger context of innovation policy.

The Article proceeds in three parts. Part I discusses the Federal Circuit's assertion of power over factual issues and the manner in which this arrogation conflicts with existing principles of power allocation between administrative agencies, trial courts, and appellate courts. Part II argues that although the PTO and trial courts have significant gaps in their fact-finding expertise, *de novo* review of technological fact and formalist rules that eliminate a role for factual inquiry are not the answer. This Part also outlines mechanisms to vest greater expertise in the PTO and trial courts and thus give appellate courts greater justification for relinquishing control over fact finding. Part III discusses how the Federal Circuit's formalist jurisprudence substitutes bright-line rules not only for factual inquiry but also for the policy analysis required by the patent statute. After discussing the limitations of the Federal Circuit's formalist analysis, it provides an outline of how appellate review in a reorganized patent system might function.

15. For examples of policy directives that depend on factual inquiry, see *infra* text accompanying note 302.

I. THE FEDERAL CIRCUIT AS FACT FINDER

In this Part, I discuss how the Federal Circuit has, in contravention of established doctrines of administrative law and appellate review, arrogated power over factual questions. Such a discussion presupposes, of course, that issues that are primarily factual in nature can meaningfully be distinguished from issues that are primarily legal in nature.¹⁶ Although a full discussion of the complex law/fact question is beyond the scope of this paper, I address briefly the reasons that this threshold presumption makes particular sense in the context of patent law.

A. *The Law/Fact Distinction*

In evaluating the “vexing” distinction between questions of law and questions of fact,¹⁷ some critics have concluded that there is no coherent mechanism for drawing such a distinction. On this view, “[m]atters of law grow downward into roots of fact, and matters of fact reach upward, without a break, into matters of law.”¹⁸ The law/fact distinction may also be vulnerable to attack from two more theoretical vantage points, one drawing on positivist jurisprudence, the other on postmodern philosophy of science. From the standpoint of a legal positivist, there is no clear distinction between law and fact: Law is itself “a matter of social fact in the sense that the authority of the rule of recognition is itself a matter of social convention.”¹⁹ For the postmodern philosopher of science, the blurring of distinctions comes from the opposite direction: Statements of fact, even facts about the physical world, cannot be separated from other types of statements because factual observations themselves are “theory laden” (and hence “value laden”).²⁰

Each of these critiques has some merit. By the same token, they do not eviscerate the principle that legal determinations tend to differ in systematic ways from the technical, case- and time-specific “adjudicative”

16. Although I follow the conventional terminology in discussing the difference between factual questions and legal questions, the language of the patent statute is sufficiently open-ended that it would be more accurate to contrast factual questions on the one hand with legal/policy questions on the other. For a discussion of why patent law is, or at least should be, largely patent policy, see *infra* Parts II.C.2–3.

17. *Pullman-Standard v. Swint*, 456 U.S. 273, 288 (1982); see also Gary Lawson, *Proving the Law*, 86 Nw. U. L. Rev. 859, 863 (1992) (noting that the law/fact distinction is “sometimes criticized as confusing and unhelpful”).

18. John Dickinson, *Administrative Justice and the Supremacy of Law in the United States* 55 (1927).

19. Jules L. Coleman, *Rules and Social Facts*, 14 Harv. J.L. & Pub. Pol’y 703, 716–17 (1991); see also Lawson, *supra* note 17, at 864–65 (pointing out that a legal naturalist who believes that moral propositions can be judged true or false would also agree that propositions of law are factual statements).

20. See, e.g., Hilary Putnam, *Reason, Truth, and History* 143–44 (1981) (discussing non-objective nature of physical facts); *Readings in the Philosophy of Science* (Baruch A. Brody & Richard E. Grandy eds., 2d ed. 1989) (collecting writings by inter alia Thomas Kuhn and Paul Feyerabend).

findings that are the foundation of many patent cases.²¹ As discussed in Part I.B below, adjudicative facts relevant to patent cases include findings regarding how the patented invention relates to the allegedly infringing invention and to the technology in the area at a particular point in time.

The first critique, which notes the manner in which questions of law tend to bleed over into questions of fact, unrealistically suggests that “law” and “fact” must either be two entirely distinct categories or be entirely indistinguishable.²² It also ignores the existence of a third category of questions, which involve the application of law to fact. Difficult classification questions often arise because scholars try to pigeonhole determinations into the categories of “pure” law or “pure” fact.²³ As many commentators have noted, inquiries like negligence fall into a third category—the category of law application (also known as the category of mixed question of law and fact).²⁴ Thus, for example, the question of where the defendant’s vehicle was located relative to an intersection when the light turned red is one of fact. Like all questions of fact, this question can be answered “without *significantly* implicating the governing legal principles.”²⁵ The question of what duty a defendant owes a plaintiff when operating her vehicle on a public street is one of law. This question requires knowledge of the applicable legal principles. After the facts are found and the law declared, the law application step necessary to determine negligence can occur.²⁶

When commentators complain about the difficulty of distinguishing law from fact—and about deciding which institution should take responsibility for questions that involve both law and fact—they are typically talking about questions of law application.²⁷ Even with law application, however, either law or fact typically predominates. Consider the follow-

21. See 2 Kenneth Culp Davis, *Administrative Law Treatise* § 12:3, at 413 (2d ed. 1979) (discussing “adjudicative” facts “pertaining to the parties and their businesses and activities”). Unless otherwise noted, my use of the term “fact” in this Article refers to adjudicative facts.

22. See Henry P. Monaghan, *Constitutional Fact Review*, 85 *Colum. L. Rev.* 229, 233 (1985) (“The incoherence argument seems greatly overdrawn once it is recognized that any distinction between ‘law’ and ‘fact’ does not imply the existence of static, polar opposites.”).

23. See Henry M. Hart, Jr. & Albert M. Sacks, *The Legal Process* 351–52 (William N. Eskridge, Jr., & Philip P. Frickey eds., 1994) (1958) (noting that confusion results from efforts to classify law application on questions as either questions of law or questions of fact, “thereby trying to squeeze into two categories questions which call for a three-fold analysis”). The desire to pigeonhole may emerge from the fact that our legal system allocates power between trial and appellate courts (and between judges and juries) according to this dichotomy.

24. In this Article, I use the terms “law application” and “mixed question of law and fact” interchangeably.

25. Monaghan, *supra* note 22, at 235. Of course, the issue of whether the evidence is legally sufficient to support a particular factual conclusion is one of law. *Id.* at 236.

26. See Hart & Sacks, *supra* note 23, at 351.

27. The problem of which institution should have primary responsibility for law application arises with particular force in the constitutional context. See Monaghan, *supra*

ing alternatives. If the legal principle is relatively specific—for example, “entering an intersection after the light has turned red constitutes negligence”—law application is a ministerial task.²⁸ What is crucial is the initial fact finding. In contrast, if the legal principle is more general—for example, “all unreasonable behavior constitutes negligence”—law application may require further elaboration of the legal principle. This elaboration will be distinctly “law-like” to the extent that it can produce rules that are properly capable of general application.²⁹ If we imagine a spectrum ranging from pure fact and ministerial law application on the one hand to law elaboration and pure law on the other, the Federal Circuit decisions that this Part discusses all involve aggressive appellate review on the fact side of this spectrum.

As for more theoretical critiques, neither the positivist critiques nor postmodern approaches undermine the important distinction that our legal system draws between law and adjudicative facts. Even if one accepts the positivist view that declarations of law represent social facts, such declarations differ significantly from adjudicative facts. Unlike adjudicative facts, legal “facts” are phrased so as to have—and indeed do have—broad applicability and effect beyond the case at hand. By contrast, findings of adjudicative fact do not (or at least should not) control future proceedings.³⁰ This is particularly true of the findings of adjudicative fact made in the patent law context: As discussed in the next section, the patent statute requires the court to make scientific and technological findings linked to a particular historical point in time (typically the time when the relevant invention was made or when the patent application was filed), with the *explicit expectation* that such findings may be different at a different point in time. Similarly, even if one accepts the postmodern view that findings of adjudicative fact may be value-laden, such findings do not have any sort of prospective force. To the contrary, the patent statute explicitly provides for factual change as a function of time.

B. *The Role of Facts in Patent Law*

In order to understand the crucial role of facts in patent law, it is first necessary to understand the law’s foundational principles. A patent is a written document that describes and claims an invention much like a land deed might describe and claim a piece of property. The patent begins with a detailed description of the invention known as the specification. This specification is followed by patent claims which define the scope of the invention—in other words the precise “metes and bounds”

note 22, at 267–69 (discussing difficulty in determining the proper role of the Supreme Court when review involves questions that require constitutional law application).

28. See *id.* at 236 (“If all legal propositions could be formulated in great detail, [law application] would be rather mechanical and require no distinctive consideration.”).

29. See *id.* at 236–37.

30. See Hart & Sacks, *supra* note 23, at 350.

of the inventive territory the patent purports to cover.³¹ Thus, in order to define patent scope, the decisionmaker must interpret the claims, an exercise known in patent law as claim construction.

When the PTO examines a patent application in order to determine whether to grant it, it first construes the claims to determine exactly how much inventive territory the applicant is claiming. The patent examiner then looks to see whether the invention as claimed satisfies a number of basic requirements contained in the patent statute. For example, the claimed invention, which can be a product or a process, must be both useful³² and novel.³³ The patent applicant must also show that she has disclosed enough information in the patent to enable others to make and use her claimed invention.³⁴ Finally, and perhaps most importantly, the patent applicant's invention must be "nonobvious"³⁵—roughly speaking, it must represent an inventive step beyond what has come before. Usefulness, novelty, adequate disclosure, and nonobviousness are all required for a patent to be valid. During the PTO examination process, patent claims often have to be narrowed in order for the PTO to be satisfied that the applicant has met all the criteria for validity.

Scope and validity are often tested once again when the holder of the patent brings an allegation of patent infringement (or a possible infringer brings a declaratory judgment action challenging patent validity). Such cases are brought in federal district court.³⁶ As noted, patent scope is determined through the exercise of claim construction. If, given the claimed scope, the trial court determines that the patent is valid, the court can then move to the issue of infringement. (Conversely, if the patent is invalid, there is no infringement issue to be decided.) Resolving the infringement question also requires looking at patent scope. The court must determine whether the defendant's invention falls within the "metes and bounds" of the patent's claims.

Although facts play a central role in many areas of law,³⁷ adjudicative facts are particularly critical to patent law. The two central determinations in patent law, those involving patent scope and patent validity, require the application of law to a set of factual findings. Depending on the specific inquiry involved, we may be concerned with facts that existed at the time the patent application was filed or facts that existed when the

31. See *Markman v. Westview Instruments, Inc.*, 517 U.S. 370, 373–74 (1996).

32. See 35 U.S.C. § 101 (2000) (containing usefulness requirement).

33. See *id.* § 102 (containing novelty requirement).

34. See *id.* § 112 (containing disclosure requirement).

35. *Id.* § 103.

36. *Id.* § 141.

37. See, e.g., Stuart Minor Benjamin, *Stepping Into the Same River Twice: Rapidly Changing Facts and the Appellate Process*, 78 Tex. L. Rev. 269, 273–75 (1999) (discussing the centrality of both adjudicative and legislative facts to constitutional jurisprudence); Kenneth Culp Davis, *An Approach to Problems of Evidence in the Administrative Process*, 55 Harv. L. Rev. 364, 402 (1942) (noting that the application of legal tests "must depend on fact-finding").

invention was made. The focus on factual findings specific to a particular point in time arises because the patent system attempts to map in a coherent way onto the scientific innovation for which it provides incentives. The underlying policy presumption is a reasonable one: If we are going to provide incentives for innovation, the incentives should track, as accurately as possible, the ever-shifting reality of the scientific and technical challenges faced by innovators.³⁸ The mechanism that the patent statute uses to perform this mapping onto the “real world” of innovation is a construct known as the “person having ordinary skill in the art”³⁹ (PHOSITA). As discussed further below, the PHOSITA is central to determinations of both patent scope and patent validity.

1. *The Role of Facts in Patent Scope.* — As noted above, patent scope is determined by construing the claims listed in the patent text. Many texts with which judges must deal, particularly statutory texts, are directed at the ordinary speaker of English. All judges are presumably ordinary speakers of English and thus are qualified to interpret these texts according to their ordinary meaning.⁴⁰ By contrast, under long-established patent case law, patent claims are *not* directed at the ordinary speaker of English; rather, they are directed at the aforementioned PHOSITA. Claim construction therefore requires a determination of what the language of the patent claims would have meant to the PHOSITA at the time the invention was made.⁴¹ Because the typical judge is not likely to be a person of ordinary skill in the relevant scientific or technological art, she is not likely to be endowed with the appropriate technical knowledge. As a consequence, even after examining a claim term in light of the accompanying specification⁴² (or, more generally, in light of any of the applica-

38. Of course, to the extent one views patent rights as similar to rights in tangible property, one could argue that initial allocations of patent rights do not matter because, absent transaction costs, parties will transfer the rights to their highest-valued uses. See generally R. H. Coase, *The Problem of Social Cost*, 3 J.L. & Econ. 1 (1960). On this view, there would presumably be little need for government institutions to perform fact-intensive inquiries in allocating rights. Even theories of patent law that analogize patent rights to rights in tangible property appear to recognize, however, that the nature of the initial allocation does matter. See *infra* notes 174, 353–355 and accompanying text.

39. 35 U.S.C. § 103 (2000).

40. See, e.g., *Green v. Bock Laundry Mach. Co.*, 490 U.S. 504, 528 (1989) (arguing that the construction of statutory terms should be based on “which meaning is . . . most in accord with context and ordinary usage”). Ordinary meaning is, of course, particularly central to textualist theories of statutory interpretation. But other theories of statutory interpretation also pay attention to the ordinary meaning of the text.

41. *Markman v. Westview Instruments, Inc.*, 52 F.3d 967, 986 (Fed. Cir. 1995) (en banc) (noting that a claim term means “what one of ordinary skill in the art at the time of the invention would have understood the term to mean”).

42. See, e.g., *Vitronics Corp. v. Conceptronic, Inc.*, 90 F.3d 1576, 1582 (Fed. Cir. 1996) (noting rule that patent claim should be interpreted in light of specification). The Federal Circuit has also held that claim terms should be construed in light of statements made by the patentee during the prosecution of the patent before the PTO. See, e.g., *Rexnord Corp. v. Laitram Corp.*, 274 F.3d 1336, 1343 (Fed. Cir. 2001). The Federal Circuit’s reliance on this so-called prosecution history has, however, been controversial.

ble canons of claim construction)⁴³ she might still find the term difficult to understand. Indeed, to the extent that the judge assumes that she knows the meaning of a particular term, it may be that she is making incorrect assumptions about how one skilled in the art would interpret the language.

In many cases involving technically complex invention, the judge would be well advised to turn to external factual evidence (known in the patent lexicon as “extrinsic evidence”) on how the term is interpreted in the relevant scientific or technological community. The famous case of *Markman v. Westview Instruments* illustrates the centrality of fact. The Federal Circuit opinion in that case reached two conclusions: first, that claim construction was solely an issue of law to be reviewed de novo; and second, that claim construction should be decided by the judge, not the jury.⁴⁴ The Supreme Court opinion squarely addressed only the second question, on which it affirmed the Federal Circuit.⁴⁵ Nonetheless, the Court went out of its way to note that discerning the meaning of a claim term within a trade or profession can be an “evidentiary” investigation.⁴⁶ Based on this recognition, the Court observed that claim construction

See John R. Thomas, On Preparatory Texts and Proprietary Technologies, 47 UCLA L. Rev. 183, 186–87 (1999) (noting that prosecution history records are often uneven and inaccurate).

43. Like canons of statutory construction, canons of claim construction assist the court in interpreting language consistently. By the same token, like canons of statutory construction, these canons can sometimes be in tension with each other. As noted in the text, claim language is supposed to be interpreted in light of the invention embodiments discussed in the specification. However, it is not supposed to be limited to the particular embodiments of the invention discussed in the specification. See *Vitronics*, 90 F.3d at 1582. The obvious tensions between the various canons of claim construction underscore the fact that the judge may frequently need to move beyond canons. Indeed, in some cases, even the same canon may lead to varying interpretations. See, e.g., *infra* text accompanying notes 71–73 (discussing *In re Baker Hughes, Inc.*, 215 F.3d 1297 (Fed. Cir. 2000)).

44. *Markman*, 52 F.3d at 967.

45. *Markman v. Westview Instruments, Inc.*, 517 U.S. 370, 372 (1996).

46. *Id.* at 390–91. In some respects claim construction is analogous to the interpretation of contractual terms. Courts often look at factual evidence when determining the meaning of an ambiguous contract term. This evidence may include “all writings, oral statements, and other conduct by which the parties manifested their assent, together with any prior negotiations between them and any applicable course of dealing, course of performance, or usage.” See E. Allan Farnsworth, *Contracts* § 7.10, at 511 (2d ed. 1990). Claim construction differs from contract interpretation to the extent that the subjective intent of the patentee does not have weight in claim construction. The Federal Circuit has analogized claim construction not to contract interpretation but to statutory interpretation. Under most theories of statutory interpretation, however, language is to be interpreted (at least in part) from the standpoint of the ordinary speaker of English. By contrast, ordinary meaning of this sort has limited value in claim construction: The relevant audience is not the ordinary speaker of English, but a person of skill in the relevant art. Cf. Craig Allen Nard, *A Theory of Claim Interpretation*, 14 *Harvard J.L. & Tech.* 1, 4–5 (2000) (accepting analogy between claim construction and statutory interpretation but arguing that the Federal Circuit has adopted a “hypertextualist” view of claim construction).

was a “mongrel practice” that “falls somewhere between a pristine legal standard and a simple historical fact.”⁴⁷ On this issue, trial judges overwhelmingly agree with the Supreme Court. A 1999 survey by the American Bar Association (ABA) Section of Intellectual Property Law found that trial courts considered extrinsic evidence in eighty-three percent of cases that required claim construction.⁴⁸

Because the Supreme Court’s observations regarding the role of fact in claim construction were not necessary to its disposition of the *Markman* case, the Federal Circuit has discounted these observations. But even decisions by the Federal Circuit that attempt to deny any significant role for facts in claim construction⁴⁹ unwittingly reveal the importance of fact. Indeed, the Federal Circuit’s own decision in *Markman* acknowledged that, because claims are directed to those skilled in the art, evidence such as expert testimony might assist the court in claim construction.⁵⁰ Nonetheless, it argued that sorting this evidence by “using certain . . . evidence that the court finds helpful and rejecting other evidence as unhelpful” did not call for fact finding because the court would not be “crediting certain evidence over other evidence or making factual evidentiary findings.”⁵¹ How accepting certain evidence as helpful or rejecting other evi-

47. *Markman*, 517 U.S. at 378, 388 (internal citations omitted). To be sure, the *Markman* Court did affirm the Federal Circuit on the specific question it took for review: whether the task of claim construction should be performed by the judge or by the jury. The Court did so, however, on the basis of largely functional considerations, deciding that such interpretation could best be done by a judge. *Id.* at 388 (“Where history and precedent provide no clear answers, functional considerations also play their part in the choice between judge and jury to define terms of art.”). I discuss the Court’s somewhat surprising emphasis on functional considerations in deciding the Seventh Amendment issue *infra* notes 266–275 and accompanying text.

48. See 1999 ABA Section of Intellectual Property Law, 1999 *Markman* Survey, 18 A.B.A. Sec. Pub. I.P.L. 3 (2000) [hereinafter *Markman* Survey].

49. See, e.g., *Cybor Corp. v. FAS Techs.*, 138 F.3d 1448, 1455 (Fed. Cir. 1998) (en banc) (“[C]laim construction is a pure issue of law”); *Markman*, 52 F.3d at 977–79 (“[C]ourts should construe patent claims as a matter of law and should not give such tasks to the jury as a factual matter[.] It has long been and continues to be a fundamental principle of American law that the construction of a written evidence is exclusively with the court.”). The later en banc decision, *Cybor*, was necessary because the Supreme Court’s disposition of the *Markman* case noted (though it did not specifically hold) that claim construction had factual underpinnings. The Supreme Court’s pronouncement led to renewed conflict between those judges on the Federal Circuit who regarded claim construction as having factual underpinnings and those who thought of it as strictly a question of law. In particular, opinions by Judges Rader and Mayer emphasized that claim construction was a mixed question of law and fact and that the trial court’s findings of fact should not be set aside absent clear error. See *Eastman Kodak Co. v. Goodyear Tire & Rubber Co.*, 114 F.3d 1547, 1555–56 (Fed. Cir. 1997) (Rader, J.); *Metalllics Sys. Co., L.P. v. Cooper*, 100 F.3d 938, 939 (Fed. Cir. 1996). *Cybor* specifically disavowed these opinions.

50. *Markman*, 52 F.3d at 980 (“The court may, in its discretion, receive extrinsic evidence in order to aid the court in coming to a correct conclusion as to the true meaning of the language employed” (internal quotes and citations omitted)).

51. *Id.* at 981.

dence as unhelpful differs from “crediting certain evidence over other evidence” is not entirely clear.⁵²

In some recent panel opinions, the Federal Circuit has attempted to soften the logical contradiction of a position that concedes a role for factual evidence while simultaneously denying any role for fact finding. This effort has had dubious success. For example, in *Vitronics Corp. v. Conceptronic, Inc.*, the Federal Circuit stated that the use of extrinsic evidence in claim construction is “rarely, if ever,” proper⁵³—a statement that flies in the face of the principle that claims should be interpreted from the standpoint of one who is skilled in the art. Indeed, in *Pitney-Bowes v. Hewlett-Packard*, Judge Michel, the author of the *Vitronics* opinion, pointedly retreated from his prior rejection of extrinsic evidence.⁵⁴ He noted that a trial court’s decision to consult “trustworthy” extrinsic evidence is “entirely appropriate, perhaps even preferable” for the purposes of ensuring that “the claim construction it is tending to from the patent file is not inconsistent with clearly expressed, plainly apposite, and widely held understandings in the pertinent technical field.”⁵⁵ Once again, though, how the Federal Circuit proposes to distinguish the determination that evidence is “trustworthy” from fact finding is not at all clear.

2. *The Role of Facts in Patent Validity: Nonobviousness and Disclosure.* — Just as claim construction involves fact finding, so too does patent validity. The central questions in the determination of patent validity are the inquiries regarding nonobviousness and adequate disclosure. In what follows, I consider the role of facts in each inquiry.

The language of the patent statute requires the patentee to show that, at the time of its invention, the subject matter of her patent application would not have been obvious to the PHOSITA.⁵⁶ Although courts

52. See *id.* at 1002–08 (Newman, J. dissenting). Similarly, Judge Rader has noted that the Federal Circuit’s instruction that trial judges should use extrinsic evidence “to understand, but not to interpret, the claim terms . . . is difficult to grasp.” See *Cybor*, 138 F.3d at 1474 (Rader, J., dissenting in part and concurring in part).

53. 90 F.3d 1576, 1585 (Fed. Cir. 1996).

54. 182 F.3d 1298 (Fed. Cir. 1999).

55. *Id.* at 1309.

56. 35 U.S.C. § 103 reads in part:

A patent may not be obtained . . . if the differences between the subject matter to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains.

35 U.S.C. § 103(a) (2000). Empirical work by Mark Lemley and John Allison indicates that nonobviousness is the most important criterion in determining patent validity; forty-two percent of patents that are held invalid in litigation are invalidated on grounds of nonobviousness. See John R. Allison & Mark A. Lemley, *Empirical Evidence on the Validity of Litigated Patents*, 26 *AIPLA Q.J.* 185, 208 (1998). Another important (and somewhat related) ground for finding patents invalid is lack of novelty under § 102. *Id.* (noting that 26.8% of patents are invalidated on grounds of lack of novelty). In these cases, the invention is not merely nonobvious given the prior art but is actually already found in the prior art. As with nonobviousness, facts are central to the novelty determination. Even the Federal Circuit has recognized this reality. See *Rapoport v.*

have specified factors that go into determining nonobviousness—and, thus, it is generally deemed an inquiry that requires the application of law to fact⁵⁷—the inquiry in most cases is highly fact-specific. As the Supreme Court has emphasized, the inquiry is necessarily based on factual questions regarding the “scope and content” of the prior invention (known as “prior art”) in the field at the time of invention; “differences between the prior art and the claims at issue”; and “the level of the ordinary skill in the pertinent art.”⁵⁸ So-called “secondary considerations,” which can help prove nonobviousness (such as the commercial success of an invention or a “long-felt but unsolved” need for the invention in the relevant art), are also factual determinations.⁵⁹ In many if not most cases, experts can testify as to these factual questions. Finally, a specific issue regarding the scope and content of the prior art that often arises in non-obviousness cases is the extent to which prior art references can be combined so as to defeat a claim of nonobviousness. Because this issue requires an inquiry into whether the particular prior art references themselves suggested combination, or whether the desire to combine would have been evident to a person of ordinary skill, it is best characterized as a question of fact.⁶⁰

Like the nonobviousness inquiry, the inquiry regarding adequate disclosure is founded in fact. The most important component of adequate disclosure, the enablement requirement, holds that the patentee must disclose to the PHOSITA how to make and use the patented invention

Dement, 254 F.3d 1053, 1057 (Fed. Cir. 2001) (noting that anticipation, one mechanism for proving lack of novelty, is a question of fact).

57. The Supreme Court has said, for example, that the determination of nonobviousness is comparable to the determination of negligence. See *Graham v. John Deere Co.*, 383 U.S. 1, 18 (1966). Negligence is perhaps the quintessential example of law application. See *supra* text accompanying notes 24–29.

58. *Graham*, 383 U.S. at 17. Certain judges on the Federal Circuit have also recognized the fact-intensive nature of obviousness. For example, in *Litton Sys., Inc. v. Honeywell, Inc.*, Judge Rader applied the lenient “substantial evidence” standard of review to a jury determination of nonobviousness by emphasizing that it is a “highly fact-specific inquiry.” 87 F.3d 1559, 1570 (Fed. Cir. 1996). However, the Federal Circuit has, in a large number of cases (particularly cases on direct appeal from the PTO), exercised plenary review over nonobviousness by simply asserting that it is a question of law. See *infra* text accompanying notes 74–79.

59. See *Graham*, 383 U.S. at 17–18 (discussing use of secondary considerations). Secondary considerations are not mentioned in the patent statute. The secondary considerations doctrine is, however, a clearly recognized component of patent law. The doctrine thus represents one of many examples of common law development within the patent law. See *infra* Part III.C.2.

60. Even the Federal Circuit has regarded the question of motivation to combine prior art references as one of fact. See *McGinley v. Franklin Sports, Inc.*, 262 F.3d 1339, 1351 (Fed. Cir. 2001) (noting that the motivation question can be viewed as a subset of the inquiry into the scope and content of the prior art but is also informed by the level of skill in the art and an appreciation of the differences between the patent claims at issue in the case and the prior art).

without “undue experimentation.”⁶¹ Although the ultimate conclusion regarding enablement (like that regarding nonobviousness) requires law application, the enablement inquiry requires underlying factual findings regarding the level of skill in the art.⁶² Another component of adequate disclosure that has achieved some prominence in the Federal Circuit’s recent jurisprudence is the written description requirement.⁶³ The precise contours of this requirement are far from clear⁶⁴ and, indeed, the requirement itself has been very controversial.⁶⁵ Roughly speaking, written description appears to require that the patentee show, through her specification, that she has actual possession of her invention. Possession, in turn, is a term of art indicating that, from the standpoint of the PHOSITA, the patentee has shown that she has actually made the invention. Thus it is clear that determining whether the written description requirement has been satisfied is a fact centered inquiry. The inquiry must be made at the time the description is filed, and it must be made from the standpoint of the PHOSITA.⁶⁶ Even the Federal Circuit has acknowledged that the question of whether the written description requirement has been satisfied turns largely on fact.⁶⁷

With that background on the law/fact distinction and on the important role of facts in the two central patent law inquiries of claim construction and validity, we can discuss the Federal Circuit’s review of PTO and trial court decisions. In the next section, I consider first the situation where the patent applicant, denied approval by the PTO, appeals directly to the Federal Circuit. I then discuss the Federal Circuit’s aggressive review of fact finding in trial court decisions.

61. *In re Wright*, 999 F.2d 1557, 1561 (Fed. Cir. 1993); *In re Vaeck*, 947 F.2d 488, 495 (Fed. Cir. 1991).

62. See *PPG Indus., Inc. v. Guardian Indus. Corp.*, 75 F.3d 1558, 1564 (Fed. Cir. 1996) (observing that enablement is a question of law based on underlying factual findings).

63. See, e.g., *Gentry Gallery, Inc. v. Berkline Corp.*, 134 F.3d 1473, 1479–80 (Fed. Cir. 1998) (applying written description requirement to narrow scope of patent claims on sectional sofa); *Regents of the Univ. of Cal. v. Eli Lilly & Co.*, 119 F.3d 1559, 1575 (Fed. Cir. 1997) (deeming patent relating to recombinant DNA technology invalid for failure to comply with written description requirement).

64. See, e.g., Fed. Cir. Bar Assoc. Patent and Trademark Appeals Comm., *Conflicts in Federal Circuit Patent Law Decisions*, 11 Fed. Cir. B.J. 723, 725 (2001) (listing as the leading conflict the written description requirement and noting that “the Federal Circuit has not provided clear and consistent rules for determining precisely what type of disclosure is sufficient to comply with the § 112 written description requirement”).

65. See, e.g., Mark D. Janis, *On Courts Herding Cats: Contending with the “Written Description” Requirement (and Other Unruly Patent Disclosure Doctrines)*, 2 Wash. U. J.L. & Pol’y 55, 58–62 (2000) (arguing that there is no separate written description requirement in the patent statute).

66. See *In re Alton*, 76 F.3d 1168, 1175 (Fed. Cir. 1996).

67. See, e.g., *id.* at 1174; *Vas-Cath Inc. v. Mahurkar*, 935 F.2d 1555, 1561 (Fed. Cir. 1991).

C. Federal Circuit Review of PTO Fact Finding

1. *Direct Review of PTO Patent Denials.* — The standard of review to be applied to direct appellate review of PTO fact finding should be rather straightforward. The patent statute itself does not provide a standard of review for patent denials.⁶⁸ Hence the Administrative Procedure Act is the source of the relevant standard. Under section 706 of the Administrative Procedure Act (APA), courts are supposed to review fact finding by administrative agencies using a highly deferential approach that rejects the fact finding only if it is “arbitrary or capricious” or unsupported by “substantial evidence.”⁶⁹ For most of its existence, however, the Federal Circuit has denied that it is subject to the APA standard. Rather than applying the APA standard, it has argued that it should apply the less deferential “clearly erroneous” standard that appellate courts apply to district courts. In various important cases, moreover, even the clearly erroneous standard has been nowhere in evidence. The Supreme Court’s recent decision in *Dickinson v. Zurko*,⁷⁰ discussed further below, may have some impact on Federal Circuit review practice with respect to questions the court explicitly recognizes are factual in nature. But it will have no impact on review in the many cases, such as those discussed below, in which the Federal Circuit refuses to recognize the existence of factual disputes.

a. *Patent Scope.* — An easy context in which to observe the absence of deferential view is claim construction. Because the Federal Circuit has (contrary to the Supreme Court’s observations in *Markman*) defined claim construction as a question of pure law, there is, in the Federal Circuit’s estimation, no need for deference.⁷¹ Nonetheless, the Federal Circuit’s de novo review of PTO claim construction in the context of certain patent denials unwittingly demonstrates the serious indeterminacy that can arise when claim construction is done by applying a pure legal analytic framework. In one recent case, *In re Baker Hughes, Inc.*, which involved methods for limiting the escape of hydrogen sulfide gas from hy-

68. As discussed *infra* text accompanying notes 91–92, the patent statute’s enunciation of a standard of review is limited to patent grants.

69. Indeed, in cases decided prior to the enactment of the APA, the Supreme Court suggested that in cases involving law application, courts should defer not only to the agency’s factual findings but also to its ultimate determination of the application question. See *NLRB v. Hearst Publ’ns, Inc.*, 322 U.S. 111, 131 (1944) (holding that where “the question is one of specific application of a broad statutory term in a proceeding in which the agency administering the statute must determine it initially, the reviewing court’s function is limited [to determining whether the agency’s decision has] a reasonable basis in law”).

70. 527 U.S. 150 (1999), discussed *infra* text accompanying notes 79–89.

71. See, e.g., *In re Baker Hughes, Inc.*, 215 F.3d 1297, 1300 (Fed. Cir. 2000) (noting de novo review of PTO claim construction). Notably, prior to its decision in *Markman*, the Federal Circuit had said that PTO claim construction was a question of law to be reviewed de novo only if there were no underlying factual issues. See *In re Donaldson Co.*, 16 F.3d 1189, 1192 (Fed. Cir. 1994).

drocarbons, the Federal Circuit overturned a PTO patent denial that had been based on the PTO's interpretation of the claim term "hydrocarbon."⁷² According to the Federal Circuit, reading the term in light of the accompanying specification—which discussed liquid hydrocarbons in certain contexts—made it clear that the patentee's use of the term was limited to liquid hydrocarbons.⁷³ As the Federal Circuit acknowledged, however, the specification referred in other contexts to *both* liquid and gaseous hydrocarbons. Indeed, based in part on the same specification, the PTO had concluded that the patentee's use of the term encompassed both liquid and gaseous hydrocarbons. In this case, as in many others, some discussion of factual evidence regarding how the term was generally used in the relevant art might have been helpful. Not surprisingly, however, the Federal Circuit's analysis contained no such discussion.

b. *Patent Validity*. — Even in validity inquiries involving nonobviousness and enablement—determinations that the Federal Circuit itself has acknowledged have factual foundations—fact-oriented standards of review are often absent. A prominent illustration of this absence is the important biotechnology case of *In re Deuel*.⁷⁴ *Deuel* addressed the issue of whether a claim to a DNA sequence can be obvious when the method for isolating the sequence is well known in the prior art. In that case, the Federal Circuit briefly noted the factual underpinnings of nonobviousness.⁷⁵ Nonetheless, it independently reviewed, and overturned, the PTO's factual determinations regarding the scope and content of the prior art. In *Deuel*, one important factual question regarding the scope of the prior art was whether methods for isolating DNA sequences are relevant prior art for patents that claim DNA sequences. The PTO said "yes" and further noted that these methods for isolating the patent applicant's claimed sequences were, at the time of the invention, well known to, and routinely used by, scientists in the field. On those grounds, the PTO denied the DNA sequence application.⁷⁶ The Federal Circuit overturned the PTO's determination, asserting (without any hint that the PTO had conducted fact finding deserving of some level of deference) that methods should not be considered appropriate prior art for claims to DNA sequence structures.⁷⁷ In another part of the *Deuel* opinion, the Federal Circuit did entertain briefly the idea that methods might be appropriate prior art for DNA sequence claims. However, it then proceeded to reject

72. 215 F.3d at 1297.

73. *Id.* at 1299.

74. 51 F.3d 1552 (Fed. Cir. 1995).

75. *Id.* at 1558 (noting that although nonobviousness is ultimately a question of law to be reviewed *de novo*, factual findings underlie the nonobviousness determination).

76. See *Ex parte Deuel*, 33 U.S.P.Q.2d (BNA) 1445, 1447 (Bd. Pat. App. & Int. Nov. 30, 1993) (noting views of PTO patent examiners that where method for isolating a sequence is "routine," the sequence should be considered to be in the public domain).

77. See *Deuel*, 51 F.3d at 1559 (overturning PTO determination, arguing that the "PTO's focus on known methods for potentially isolating the claimed DNA molecules is . . . misplaced because the claims at issue define compounds, not methods").

the PTO's factual finding that the prior art method in question actually could be used to find the DNA sequence of interest.⁷⁸ Once again, the Federal Circuit's rejection of the PTO failed to mention the factual nature of determinations regarding relevant prior art.

Deuel is hardly an isolated case. To the contrary, as discussed further in Part II, because the Federal Circuit appears to have determined that its conclusion on the factual question of whether methods can be prior art for DNA sequence claims is actually a generalizable *legal* principle that can be cited as precedent in future cases, it is now significantly easier to meet the nonobviousness requirement in the entire area of genotechnology. In addition, as discussed below, this newfound legal principle has also transformed disclosure doctrine in the area of genotechnology.⁷⁹

c. *The Impact of Dickinson v. Zurko*. — The Supreme Court's recent decision in *Dickinson v. Zurko*⁸⁰ has challenged the Federal Circuit's refusal to show APA-level deference. In *Zurko*, a case involving the Federal Circuit's en banc reversal of PTO factual findings regarding the obviousness of a computer security system invention, the Supreme Court held that the highly deferential APA standard of review squarely applied to all Federal Circuit review of PTO fact finding. Writing for the Court, Justice Breyer noted that the language of the Administrative Procedure Act (APA) did not specifically exempt any agencies from the Act's purview.⁸¹ Thus, if the Federal Circuit's review of the PTO were to be exempt from the ordinary APA standard, its review would have to fall under a general exception to ordinary review specified in the APA. The Court rejected the contention that the Federal Circuit's non-APA standard of review fell under section 559 of the APA, which states that the Act does not "limit or repeal additional requirements . . . recognized by law." According to the Court, there was insufficient evidence that, at the time of the APA's adoption in 1946, courts had used the Federal Circuit's preferred clearly erroneous standard in reviewing PTO action.⁸² To the contrary, a pre-APA Supreme Court case that had addressed the question of judicial review of the PTO, *Morgan v. Daniels*, had indicated that the PTO's factual determinations were entitled to substantial deference as the "action of one of the executive departments of the government."⁸³

78. *Id.* at 1558–59.

79. See Part I.D.2.

80. 527 U.S. 150 (1999).

81. *Id.* at 154.

82. *Id.* at 154–55.

83. 153 U.S. 120, 124 (1894). Orin Kerr takes issue with the *Zurko* decision, arguing that because the PTO is a "non-regulatory" agency, the APA does not apply to its fact finding (or any of its actions, for that matter). See Orin S. Kerr, Rethinking Patent Law in the Administrative State, 42 Wm. & Mary L. Rev. 127, 173–77 (2000). Kerr's assertion appears to lack doctrinal foundation. Although a predecessor bill to the APA had exempted the PTO from its purview, the APA itself contains no such exemption. Even the Federal Circuit opinion in *Zurko*, which argued for the clearly erroneous standard of

The APA standard is considered more deferential than the clearly erroneous standard. Thus, to the extent that the Federal Circuit recognizes the existence of factual disputes that are subject to the Supreme Court's instruction in *Zurko*, it may defer more frequently than it has through its haphazard application of the clearly erroneous standard. It bears mention, however, that the Federal Circuit response to the Supreme Court's admonition betrays some signs of continued resistance to deference. In *Zurko*, the Supreme Court did not specifically resolve the question of whether the APA's arbitrary or capricious standard or its substantial evidence standard applied to factual findings.⁸⁴ Instead, the Court cited a D.C. Circuit opinion by then-Judge Scalia that found no difference between the level of deference entailed in substantial evidence and arbitrary or capricious standards as applied to judicial review of agency fact finding.⁸⁵ In a subsequent case interpreting *Zurko*, *In re Gartside*, the Federal Circuit determined that substantial evidence review was in fact the appropriate standard.⁸⁶ Additionally, in contrast to the Supreme Court, the Federal Circuit strenuously argued that substantial evidence review was significantly less deferential than arbitrary or capricious

review (and was therefore reversed by the Supreme Court), acknowledged that "Congress drafted the APA to apply to agencies generally" *In re Zurko*, 142 F.3d 1447, 1452 (Fed. Cir. 1998) (en banc). Moreover, the parties in the *Zurko* Supreme Court appeal agreed that the PTO was an agency subject to the APA's constraints and that the PTO's fact finding constituted "agency action." *Zurko*, 527 U.S. at 154. Kerr may be making a normative argument that because the PTO does not utilize some of the procedures of full-fledged agencies, such as adversarial proceedings, its fact finding should not be given APA-level deference, no matter how clear the plain language of the APA. Although the PTO's procedures for evaluating patent applications should clearly be augmented, see *infra* Part II.A.2.b, the deficiencies of the PTO procedure are likely to be most relevant not in the context of patent denials that are subject to the APA but in the context of patent grants. In addition, Kerr's thesis that the patent system is, or at least should be, a "private law" system that has only a nominal role for the PTO is misplaced. The Constitutional provision authorizing (but not requiring) Congress to grant patents for "limited Times" in order "to promote the Progress of . . . useful Arts" provides substantial evidence that patents have, from the very outset, been publicly administered rights that serve an instrumental social purpose. See U.S. Const. art. I, § 8, cl. 8. Indeed, Thomas Jefferson, the primary architect of the constitutional framework for patent protection and the first Patent Commissioner, emphasized in his writings that inventions

cannot, in nature, be a subject of property. Society may give an exclusive right to the profits arising from them, as an encouragement to men to pursue ideas which may produce utility, but this may or may not be done, according to the will and convenience of the society, without claim or complaint from any body.

Letter from Thomas Jefferson to Isaac McPherson (Aug. 13, 1813), in 6 *The Writings of Thomas Jefferson* 175, 180–81 (H.A. Washington, ed., New York, Riker Thorne & Co. 1854), quoted in *Graham v. John Deere Co.*, 383 U.S. 1, 9 n.2 (1966). Similarly, the Supreme Court has emphasized that "[t]he patent monopoly was not designed to secure to the inventor his natural right in his discoveries. Rather, it was a reward, an inducement, to bring forth new knowledge." *Graham*, 385 U.S. at 9.

84. 5 U.S.C. § 706(2)(A), (E) (2000).

85. *Zurko*, 527 U.S. at 158 (citing *Ass'n of Data Processing Serv. Orgs. Inc. v. Bd. of Governors of Fed. Reserve Sys.*, 745 F.2d 677, 683–84 (D.C. Cir. 1984)).

86. 203 F.3d 1305 (Fed. Cir. 2000).

review.⁸⁷ Another case that suggests Federal Circuit resistance to deference is the *Zurko* remand itself.⁸⁸ In the remand, the Federal Circuit proceeded to apply the substantial evidence standard it had adopted in *Gartside*. Even though the Federal Circuit's previous en banc opinion in the *Zurko* case had said that the choice between the clearly erroneous standard of review and the APA standard of review would be outcome-determinative on the ultimate question of nonobviousness,⁸⁹ the court once again rejected the PTO's factual findings on obviousness.⁹⁰ In other words, it would appear that the Federal Circuit refused to apply the APA standard of review even in the very case where the Supreme Court had mandated such review.

2. *Review of PTO Patent Grants.* — Direct appeal from a PTO patent denial is not the only context in which the Federal Circuit reviews PTO fact finding. Another context is in its review of PTO patent grants. The Federal Circuit's review of PTO fact finding in the context of a patent grant is quite indirect, however. While the patent statute provides for direct review by the Federal Circuit of patent denials, the Federal Circuit reviews patent grants only when they are the subject of infringement or declaratory judgment actions. Thus, the findings of fact that are squarely reviewed by the Federal Circuit are not those of the PTO, but rather, those of the district court that adjudicates the infringement or declaratory judgment action.

PTO fact finding is nonetheless important to the extent that the patent statute mandates that an issued patent (and presumably, any fact finding underlying the issued patent) enjoys a presumption of validity.⁹¹ Moreover, according to the Federal Circuit, courts should require "clear and convincing evidence" before the presumption of validity can be overridden.⁹²

The Federal Circuit's endorsement of the presumption of validity notwithstanding, the court has subjected trial court fact finding in infringement or declaratory judgment actions to a level of review that is contrary to traditional principles of appellate review. The next section considers the review that the Federal Circuit applies to trial court fact finding.

87. *Id.* at 1312.

88. *In re Zurko*, 258 F.3d 1379 (Fed. Cir. 2001).

89. *In re Zurko*, 142 F.3d 1447, 1449 (Fed. Cir. 1998) (en banc).

90. *Zurko*, 258 F.3d at 1384–86.

91. 35 U.S.C. § 282 (2000). While the Federal Circuit did, for a long while, succeed in questioning the applicability of the APA, see *supra* notes 69–90, it would be hard-pressed to question the applicability of the patent statute itself.

92. See *Medtronic Inc. v. Intermedics, Inc.*, 799 F.2d 734, 741 (Fed. Cir. 1986) (stating that § 282 "creates a presumption that a patent is valid and imposes the burden of proving invalidity on the challenger by 'clear and convincing evidence'").

D. Federal Circuit Review of Trial Court Fact Finding

Under traditional principles of appellate procedure and allocation of decision authority, factual findings made by trial courts are entitled to substantial deference. Federal Rule of Civil Procedure 52(a) provides that factual findings made by trial court judges are to be reviewed under the clearly erroneous standard, while factual findings made by juries are reviewed under the substantial evidence standard.⁹³ At a minimum, then, when a court is reviewing questions of law application that involve subsidiary findings of fact, the doctrinal framework suggests that the court should review the underlying factual findings deferentially.⁹⁴ Of course, to the extent that an appellate court refuses to recognize a particular question as involving a fact-finding component, it can avoid deference and adopt entirely de novo review. This has happened with claim construction and may well happen with the various inquiries into patent validity, principally nonobviousness.

1. *Review of Fact Finding in Trial Court Claim Construction.* — In the first few years after the Federal Circuit's establishment, a substantial body of its case law acknowledged the importance of facts in claim construction. For example, in the 1984 case *McGill Inc. v. John Zink Co.*, the court held that claim construction could have underlying factual issues that had to be submitted to a jury.⁹⁵ Specifically, the court acknowledged that the meaning of a term of art in a claim might be disputed, and external factual evidence might be needed to explain the meaning of the term.⁹⁶ Similarly, in *Johnston v. IVAC Corp.*, the Federal Circuit noted that disputed issues of fact could arise in claim interpretation when there was "genuine evidentiary conflict created by the underlying probative evidence pertinent to the claim's interpretation."⁹⁷ This underlying probative evidence could include conflicting technological evidence offered by expert witnesses regarding the meaning of claims as well as conflicting evidence regarding what transpired during the prosecution of the pat-

93. Fed. R. Civ. P. 52(a).

94. Indeed, in cases of law application, most circuit courts have adopted at least this level of deferential review. See John Duffy, On Improving the Legal Process of Claim Interpretation: Administrative Alternatives, 2 Wash. U. J.L. & Pol'y 109, 120–21 & n.46 (2000) (citing circuit court cases); see also *Ornelas v. United States*, 517 U.S. 690, 699 (1996) (stating in Fourth Amendment context that trial court's findings of historical fact should be reviewed for clear error). Some courts, perhaps most notably the Seventh Circuit, have gone further, arguing that deferential review on the ultimate question of law application is appropriate, as questions of law application do not tend to implicate the appellate court's responsibility for maintaining uniform law. See, e.g., *United States v. Baldwin*, 60 F.3d 363, 365 (7th Cir. 1995) ("This court . . . has moved decisively to the position that appellate review of determinations of mixed questions of fact and law should be governed by the standard of clear error, and not by the de novo standard.").

95. 736 F.2d 666, 672 (Fed. Cir. 1984).

96. *Id.*

97. 885 F.2d 1574, 1579 (Fed. Cir. 1989); see also *Moeller v. Ionetics, Inc.*, 794 F.2d 653, 657 (Fed. Cir. 1986) (vacating summary judgment where claim terms created underlying factual disputes).

ent.⁹⁸ The Federal Circuit also used fact-oriented standards of review when evaluating claim interpretation by trial court judges and juries: It used the substantial evidence standard when reviewing claim construction by juries⁹⁹ and the clearly erroneous standard when reviewing claim interpretation by judges in bench trials.¹⁰⁰ Even on those occasions where the court did not review the overall claim construction using fact-oriented standards, it used these standards to review the underlying factual findings.¹⁰¹ In recent years, by contrast, the Federal Circuit has declared claim construction to be entirely a question of law that is reviewed de novo. De novo review of claim construction was announced in 1995, in the court's en banc opinion in *Markman v. Westview Instruments, Inc.* It was then reiterated by the court in another en banc opinion, *Cybor Corp. v. FAS Technologies*. In addition, the court appears to have been quite serious in its application of de novo review. Two recent empirical studies estimate that the Federal Circuit has disagreed with lower court claim construction in more than one-third of all appealed cases.¹⁰² Even assuming that some of these claim construction modifications turned on legal principles (for example, the district court's failure to use a particular canon of claim construction properly) rather than fact, this rate of modification appears quite high.¹⁰³

98. *Tol-O-Matic, Inc. v. Proma Produkt-Und Mktg. Gesellschaft*, 945 F.2d 1546, 1550 (Fed. Cir. 1991).

99. See, e.g., *Bio-Rad Labs., Inc. v. Nicolet Instrument Corp.*, 739 F.2d 604, 614 (Fed. Cir. 1984) ("On appeal, we consider only whether reasonable jurors could have interpreted the claim in the manner presumed.").

100. See, e.g., *Perini America, Inc. v. Paper Converting Mach. Co.*, 832 F.2d 581, 586-87 (Fed. Cir. 1987) (upholding trial judge's interpretation of disputed claim terms under clearly erroneous standard).

101. See, e.g., *Tol-O-Matic*, 945 F.2d at 1552 (finding that substantial evidence supported jury's presumed findings of fact on disputed terms and prosecution history); *H.H. Robertson, Co. v. United Steel Deck, Inc.*, 820 F.2d 384, 389 (Fed. Cir. 1987) (reviewing judge's factual findings under clearly erroneous standard).

102. See Christian A. Chu, *Empirical Analysis of the Federal Circuit's Claim Construction Trends*, 16 *Berkeley Tech. L.J.* 1075, 1104 (2001) (stating that Federal Circuit modified claim construction in forty-four percent of cases); Kimberly A. Moore, *Are District Court Judges Equipped to Resolve Patent Cases?*, 15 *Harv. J.L. & Tech.* 1, 11 (2001) (giving figure of thirty-three percent). The discrepancy between the Moore and Chu findings may be a consequence of their having studied different time periods. While Moore's study encompasses cases decided between April 1996 and December 2000, Moore, *supra*, at 9, Chu's study encompasses cases decided between January 1998 and April 2000, Chu, *supra*, at 1092. As Moore notes, the Federal Circuit's reversal of lower court claim constructions increased after the 1998 decision in *Cybor*, which explicitly held that claim construction was to be reviewed de novo. Moore, *supra*, at 29. In addition, when Chu includes within his population Rule 36 summary affirmances (as does Moore in her study), he estimates that the Federal Circuit changed the lower court's claim construction in an average of thirty-five percent of cases. Chu, *supra*, at 1146.

103. Indeed, in certain cases, the Federal Circuit has not only modified the district court's claim construction, but it has adopted a claim construction not advanced by either of the parties. See, e.g., *Exxon Chem. Patents, Inc. v. Lubrizol Corp.*, 64 F.3d 1553, 1555 (Fed. Cir. 1995).

To be sure, selection effect problems suggest that these statistics on claim construction reversals should be read cautiously. The population of claim constructions that are appealed may not be representative of the overall population of claim constructions. Indeed, the well-known selection effect theory of George Priest and Benjamin Klein suggests that about fifty percent of claim constructions that are appealed *should* get modified.¹⁰⁴ According to Priest and Klein, parties appeal only those cases where they have substantially divergent expectations regarding the outcome of the case. Moreover, such divergence is most likely to arise when the outcome could readily go either way.¹⁰⁵ Applying the Priest and Klein “divergent expectations” model to the claim construction context, one might argue that parties choose to appeal only those claim constructions that represent genuinely difficult questions that could be decided either way. It bears mention, however, that the Priest and Klein model doesn’t work well in predicting overall appellate affirmance rates. Overall appellate affirmance rates in most courts, including the Federal Circuit, appear to hover around eighty percent.¹⁰⁶ In addition, the available evidence suggests that many trial claim constructions are in fact appealed. According to one survey, only twenty-nine percent of cases settle pursuant to a trial court claim construction.¹⁰⁷ Thus it would appear that many cases involving claim construction, not just those involving a close question, are probably appealed.

The Federal Circuit’s desire to control claim construction is hardly surprising. Claim construction is often determinative of all other questions in the case—as famously noted by Judge Rich, a longtime member of the Federal Circuit, “the claim is the name of the game.” Indeed, the Federal Circuit’s plenary review of claim construction has had something of a “domino effect,” leading the court to arrogate power over issues even it admits are largely factual, such as infringement.¹⁰⁸ The domino effect

104. George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. Legal Stud. 1, 4–5 (1984).

105. *Id.* at 19–20.

106. It is not clear, however, whether the eighty percent figure is correct for the Federal Circuit. For the Federal Circuit’s self reported affirmance rates, indicating a figure of about eighty percent, see Annual Report of the Director of the Administrative Office of the United States Courts, Table B-8 (1995–2000). But see Chu, *supra* note 101, at 1100 (finding 36.6% overall reversal rate for decisions, including summary affirmances, rendered between January 1, 1998 and April 30, 2000). To the extent that Chu’s data regarding overall reversal rate is correct, the Federal Circuit is being quite aggressive relative to other circuits in reversing *all* lower court decisions, not just claim constructions.

107. *Markman* Survey, *supra* note 48, at 3.

108. See, e.g., *Embrex, Inc. v. Serv. Eng’g Corp.*, 216 F.3d 1343, 1348–49 (Fed. Cir. 2000). Both infringement and infringement under the doctrine of equivalents are, according to the Federal Circuit, factual issues. The doctrine of equivalents inquiry is a judge-made equitable doctrine that allows a finding of infringement even when an accused invention does not infringe the literal claims of the patented invention. Under this doctrine, the fact finder makes a determination of whether the challenged invention performs the same function in the same way to achieve the same result as the patented

works as follows: Because claim construction bears heavily on the question of infringement, a decision to overturn the district court's claim construction often means that a new determination regarding infringement must be made. At that point, the Federal Circuit faces two choices. It can either remand to the district court for cumbersome new fact finding on the question of infringement, or it can simply determine the question of infringement itself. Even though, under the Federal Circuit's own jurisprudence, infringement turns on factual questions, the Federal Circuit is often tempted not to remand for a new—and costly—trial on infringement. Rather, the court simply declares that there can be no factual dispute with respect to infringement.¹⁰⁹ As a consequence, de novo review of claim construction effectively becomes de novo review of infringement.

In contrast to its stance on claim construction, the Federal Circuit has not declared validity determinations like nonobviousness and adequate disclosure to be pure questions of law subject to de novo review. Nonetheless, the court appears effectively to have applied de novo review to many of these determinations. Moreover, in the case of computer software (as in the cases of genotechnology discussed earlier¹¹⁰), the court appears to have elevated its factual conclusions to the level of legal precedent.

2. *Review of Fact Finding in Trial Court Nonobviousness and Disclosure Decisions.* — While we do not have quantitative data on the percentage of cases in which the Federal Circuit has reversed lower court determinations regarding nonobviousness and adequate disclosure, certain inferences can be drawn from data collected in recent studies by Kimberly Moore and Christian Chu. Based on a study of trial court and Federal Circuit decisionmaking between 1983 and 1999, Moore reports that the Federal Circuit affirmed seventy-eight percent of all decisions in which a fact finder made determinations about patent validity.¹¹¹ Because this percentage is almost precisely the same as the overall percentage of district court decisions the Federal Circuit reported as having affirmed during a similar period,¹¹² one can plausibly conclude that, even in contexts

invention. See *Graver Task & Mfg. Co. v. Linde Aire Prod. Co.*, 339 U.S. 605, 609–10 (1950). For discussion of the doctrine of equivalents as a federal common law doctrine, see *infra* notes 387–388 and accompanying text.

109. See, e.g., *Pall Corp. v. Hemasure Inc.*, 181 F.3d 1305, 1312 (Fed. Cir. 1999); *Chiuminatta Concrete Concepts, Inc. v. Cardinal Indus., Inc.*, 145 F.3d 1303, 1313 (Fed. Cir. 1998); *Exxon Chem. Patents, Inc. v. Lubrizol Corp.*, 64 F.3d 1553, 1560–62 (Fed. Cir. 1995). In their article on Federal Circuit “hyperactivity,” Rooklidge and Weil point to *Pall Corporation*, *Chiuminatta*, and *Exxon* as examples of the court's inappropriately making its own factual finding with respect to infringement. See Rooklidge & Weil, *supra* note 5, at 742–46. They fail to note that these cases are largely a result of the Federal Circuit's characterization of claim construction as a pure question of law, to be reviewed de novo.

110. See *supra* text accompanying notes 74–79.

111. Kimberly Moore, *Judges, Juries, and Patent Cases—An Empirical Peek Inside the Black Box*, 99 Mich. L. Rev. 365, 397 (2000) [hereinafter Moore, Black Box].

112. In the period from 1993 to 1998, the Federal Circuit's self-reported overall affirmance rate averaged 77.5%. *Id.*

where the Federal Circuit itself has acknowledged the need for fact finding, its review of this fact finding is no more deferential than its review of legal determinations made by the lower court.¹¹³ Christian Chu's data, based on a study of written Federal Circuit decisions between 1998 and 2000, report different absolute numbers but suggest a similar conclusion with respect to the Federal Circuit's concern for deferential review of fact. Interestingly, according to Chu's data, the Federal Circuit actually affirmed *fewer* jury trial decisions than it did summary judgments (forty-five versus fifty-two percent). Moreover, its rate of affirmance for bench trial decisions was only slightly higher than its summary judgment affirmance rate (fifty-two versus fifty-six percent).¹¹⁴ In other words, at least in the context of that subset of cases that is appealed, the Federal Circuit appears to be less deferential to jury decisionmaking than it is to strictly legal determinations made by the trial judge. As for fact finding by the trial judge, the court treats it about as deferentially as it does the legal determinations made by the trial judge.

A more qualitative review of the cases also indicates the Federal Circuit's propensity for de novo review. Indeed, this propensity became evident soon after the court's creation. In *Dennison Manufacturing Co. v. Panduit Corp.*, a 1986 case, the Supreme Court responded to an appellant's complaint that the Federal Circuit was exercising plenary power over the nonobviousness determination by vacating the Federal Circuit's reversal of the trial court on that issue.¹¹⁵ The Supreme Court then remanded the case to the Federal Circuit with specific instructions to explain why Rule 52(a) did not mandate deference to the trial court's factual determinations regarding nonobviousness.¹¹⁶ The Supreme Court was sufficiently concerned about the Federal Circuit's apparent arrogation of power over factual determinations that it chose the *Dennison* case as the first Federal Circuit case that it would review. Indeed, *Dennison* was one of only two cases that the Supreme Court reviewed in the decade following the Federal Circuit's inception.¹¹⁷

113. To be sure, it is possible that virtually all the bench and jury trial decisions were reversed on legal grounds. This seems highly unlikely, however.

114. Chu, *supra* note 102, at 1109.

115. 475 U.S. 809, 810–11 (1986).

116. *Id.* at 811.

117. The only other Federal Circuit case reviewed by the Supreme Court during this time was *Christianson v. Colt Indus. Operating Corp.*, 486 U.S. 800 (1988). Like *Dennison*, *Christianson* involved not a question of substantive patent law but, rather, an allocation of power issue. In *Christianson*, the Court held that the Federal Circuit did not have jurisdiction on appeal over a case in which the only patent law issue was a defense claim. Indeed, through most of the first two decades of the Federal Circuit's existence, the few Federal Circuit cases that the Supreme Court has reviewed have involved allocation of power issues. In addition to *Dennison* and *Christianson*, these include *Dickinson v. Zurko*, 527 U.S. 150 (1999), discussed *supra* notes 80–90 and accompanying text, and *Cardinal Chem. Co. v. Morton Int'l, Inc.*, 508 U.S. 83 (1993) (holding that the Federal Circuit may not vacate validity judgment after finding non-infringement). Very recently, however, the court has begun to review cases that involve not only allocation of power issues but also

On remand, the Federal Circuit dutifully discussed at some length the factual foundations of nonobviousness.¹¹⁸ Since *Dennison*, certain opinions, particularly by Judges Rader and Newman, have continued to invoke a fact-centered view of nonobviousness and have reviewed jury determinations of the issue under a substantial evidence standard.¹¹⁹ Various cases have also noted that, even if the jury's ultimate conclusion regarding nonobviousness is reviewed de novo, the jury's presumed resolution of the underlying factual disputes should be reviewed under a substantial evidence standard.¹²⁰

On many other occasions, however, the court has construed nonobviousness as a purely legal determination and has argued that de novo review should apply to all components of the nonobviousness determination, whether made by a trial judge in a bench trial or a jury in a jury trial. Thus, for example, in *Newell Cos. v. Kenney Manufacturing Co.*, a case decided soon after the Supreme Court remand in *Dennison*, the majority simply announced that there was no factual dispute with respect to nonobviousness and that it could therefore review the jury's determination of nonobviousness de novo.¹²¹ The majority declared that there was no issue of fact even though one of the patentee's main arguments, which the court rejected, was a factual assertion regarding the differences between a particular prior art patent and his own patent.¹²² In other cases, the court has mentioned in passing that the nonobviousness determination may involve factual findings to which it should defer. However, it has refused to show any deference. For example, in *Uniroyal, Inc. v. Rudkin-Wiley Corp.*, the trial court had found persuasive an expert witness's testimony that the invention at issue, an air-deflecting device for reducing wind resistance encountered by tractor-trailer vehicles, would have been suggested by the combination of two prior art references.¹²³ In addition, according to the trial court, other prior art in the field would have sug-

substantive patent law. The most prominent of these cases is the Supreme Court's reversal of the Federal Circuit in *Festo Corp. v. Shoketsu Kinzoku Kogyo Kabushiki Co.*, 535 U.S. 722 (2002). I discuss the implications of the Supreme Court's newfound interest in substantive patent law *infra* text accompanying notes 414–415.

118. *Panduit Corp. v. Dennison Mfg. Co.*, 810 F.2d 1561, 1566 (Fed. Cir. 1987).

119. See, e.g., *Litton Sys., Inc. v. Honeywell, Inc.*, 87 F.3d 1559, 1570 (Fed. Cir. 1996) (Rader, J.) (reinstating, over trial judge's rejection, jury verdict on grounds that substantial evidence supported jury conclusion that the patent was nonobvious); *Richardson v. Suzuki Motor Co.*, 868 F.2d 1226, 1237 (Fed. Cir. 1989) (Newman, J.) (affirming jury verdict that found patent nonobvious, hence valid, on grounds that substantial evidence supported this conclusion).

120. See, e.g., *Jurgens v. McKasy*, 927 F.2d 1552, 1557 (1991).

121. 864 F.2d 757, 762 (Fed. Cir. 1988).

122. *Id.* at 768 (rejecting patentee's argument that the prior art taught "away from a do-it-yourself adjustable shade"). Dissenting in that case, Judge Newman attacked the panel majority for baldly asserting that the facts were undisputed on appeal. *Id.* at 773 (Newman, J., dissenting).

123. See 837 F.2d 1044, 1052 (Fed. Cir. 1988).

gested combining the two references.¹²⁴ The Federal Circuit's *Uniroyal* opinion noted that factual findings underlying a nonobviousness determination should be overturned on appeal only if they are clearly erroneous. Nonetheless, the appellate court reexamined the prior art in question and summarily concluded,¹²⁵ without any apparent deference to the trial court's fact finding, that there was no basis for the conclusion that it would have been obvious to one skilled in the art to combine the prior art references.¹²⁶

Similarly, in various important cases involving the disclosure requirements, the court has merely paid lip service to deference. For example, in *Northern Telecommunications, Inc. v. Datapoint Corp.*, a case involving claims to a method for entering, verifying, and storing data using a batch data entry terminal, the Federal Circuit was faced with a decision by the lower court finding that the method claims in question were not enabled.¹²⁷ The district court had concluded that, because the patentee had not provided any details (e.g. source code or flow charts) regarding the data entry program in question, undue experimentation would be required to write the program.¹²⁸ The Federal Circuit noted that the amount of disclosure required to enable a software-based invention generally varies depending on the facts of the particular case. These facts might include "the nature of the invention, the role of the program in carrying it out, and the complexity of the contemplated programming, all from the viewpoint of the skilled programmer."¹²⁹ The Federal Circuit also acknowledged that a number of expert witnesses had testified to the effect that further detail regarding the program would indeed have been useful "in order to avoid spending experimental time."¹³⁰ The Federal Circuit even mentioned that the district court's determination regarding undue experimentation should be overturned only for clear error.¹³¹ Nonetheless, the appellate court proceeded to reverse the trial court's enablement finding, and it did so without any explanation as to why the district court's fact finding was clearly erroneous. *Northern Telecommunications* is not an isolated case in the Federal Circuit's computer software jurisprudence. To the contrary, the Federal Circuit has repeatedly enunciated a general rule that software disclosure requirements can be satisfied simply through the disclosure of software function.¹³²

124. See *id.* (citing trial court findings).

125. Its reexamination of the prior art took a total of one paragraph. See *id.*

126. *Id.*

127. 908 F.2d 931, 933 (Fed. Cir. 1990).

128. *Id.* at 943.

129. *Id.* at 941.

130. *Id.* at 942 (noting testimony of expert witness).

131. *Id.* at 943.

132. See, e.g., *Fonar Corp. v. Gen. Elec. Co.*, 107 F.3d 1543, 1549 (Fed. Cir. 1997) (arguing that "best mode" component of section 112 disclosure is satisfied so long as function is disclosed); see also *Moba, B.V. v. Diamond Automation, Inc.*, No. 01-1063, 2003 U.S. App. LEXIS 6285, at *47-*48 (Fed. Cir. Apr. 1, 2003) (Rader, J., concurring) (noting

The Federal Circuit's written description jurisprudence also demonstrates its tendency to treat fact and law interchangeably. The impetus for the rise of this jurisprudence appears to have been the *Deuel* line of non-obviousness cases discussed earlier.¹³³ Once the court had determined, as a general legal matter, that methods could not be relevant prior art for purposes of showing the obviousness of DNA sequence claims, it had to conclude, virtually as a matter of stare decisis, that such methods could not be used as adequate disclosure for such claims.¹³⁴ The written description doctrine, which had previously been used only in very limited circumstances largely unrelated to disclosure,¹³⁵ became the doctrinal box into which to fit this conclusion. Thus, in *Regents of the University of California v. Eli Lilly & Co.*, the Federal Circuit applied the written description doctrine for the first time in the ordinary disclosure context.¹³⁶ The court also announced that satisfying written description in the DNA sequence context required a showing of structure, formula, or the like; in arguing that disclosure indicative of structure or formula was necessary as a legal matter, the court was presumably declaring that no other type of disclosure could ever serve to describe the invention to the PHOSITA.¹³⁷

that since *Northern Telecommunications*, the Federal Circuit has expressly rejected a requirement of source code disclosure for software); Dan L. Burk & Mark A. Lemley, *Is Patent Law Technology-Specific?*, 17 *Berkeley Tech. L.J.* 1155, 1162 (2002) (noting that recent Federal Circuit decisions have "all but eliminated" the enablement requirement in the area of computer software).

133. See *supra* text accompanying notes 74–79.

134. See *Regents of the Univ. of Cal. v. Eli Lilly & Co.*, 119 F.3d 1559, 1567 (Fed. Cir. 1997).

135. In a recent dissent from a denial of rehearing en banc (a rehearing that would have revisited the question of whether written description is a distinct disclosure doctrine), Judge Rader noted that written description had previously been used only in the context of later filed claims to ensure that these claims did not introduce "new matter." *Enzo Biochem, Inc. v. Gen-Probe, Inc.*, 42 F. Appx. 439, 446–48 (Fed. Cir. 2002) (Rader, J., dissenting); see also Arti K. Rai, *Intellectual Property Rights in Biotechnology: Addressing New Technology*, 34 *Wake Forest L. Rev.* 827, 834 (1999) [hereinafter Rai, *Addressing New Technology*] (noting that "the [Federal Circuit] has used the written description requirement in a manner that somewhat raises the patentability bar").

136. 119 F.3d at 1567.

137. In a recent case, *Enzo Biochem, Inc. v. Gen-Probe, Inc.*, 285 F.3d 1013 (Fed. Cir. 2002), vacated by 296 F.3d 1316 (Fed. Cir. 2002), Judge Lourie, the author of the *Deuel* and *Lilly* opinions, attempted to apply his structure or formula rule without any inquiry into the underlying biological facts. Because the structure or formula rule is itself ungrounded factually, the *Enzo* decision was doubly removed from contact with factual findings. In *Enzo*, the Federal Circuit affirmed the lower court's summary judgment determination that the patent in question, which claimed nucleic acid probes, lacked written description. *Id.* at 1018–24. The district court had granted summary judgment without hearing any expert testimony on the question of whether the patentee's disclosure would have revealed nucleic acid structure. See *id.* at 1024 (Dyk, J., dissenting). The Federal Circuit similarly concluded as a matter of law that its written description rule had not been satisfied by the patentee's disclosure. *Id.* at 1018–24. The appeals court reached this conclusion as a matter of law even though there were strong arguments that, as a simple matter of biology, the patentee's revelation of hybridization function inevitably revealed at least some

Thus far I have argued that the Federal Circuit's review of PTO and trial court decisionmaking contravenes—in rather dramatic fashion—the traditional doctrinal framework for judicial review of fact finding by administrative agencies and trial courts. This traditional framework may, however, be inapplicable as a normative matter. The next Part addresses the normative question of whether, given the infirmities of the inferior decisionmakers, the Federal Circuit should continue to maintain tight control over fact finding.

II. THE NORMATIVE ANALYSIS: SHOULD THE FEDERAL CIRCUIT EXERCISE DE NOVO REVIEW?

As various commentators have noted, the division of responsibility between administrative agencies, trial courts, and appellate courts should rest (and, indeed, probably does rest) not only on whether a question is primarily legal or factual but also on functional considerations such as the relative competence of the institutions.¹³⁸ In this case, conventional review doctrine may simply not be appropriate for a specialized court such as the Federal Circuit, particularly if the inferior decisionmakers are suspect. In this Part, I address the normative justification for appellate court deference. Even under the current system, deference to fact finding that is the basis for a PTO patent denial is probably warranted. Federal Circuit deference to PTO fact finding in the context of patent *grants*, or to district court fact finding more generally, poses a more difficult question. The costs of delaying decisionmaking on basic factual questions until a patent case reaches the appellate level are extremely high, however. In addition, because the Federal Circuit often decides these factual questions incorrectly, and then gives its erroneous determinations precedential weight, the Federal Circuit's fact finding has had problematic effects across entire fields of technology. So long as the PTO and the trial courts can be restructured in a manner that produces adequate fact finding, deference should be given. In what follows, I make the case for why primary responsibility for fact finding—and for law application where the case turns on factual findings rather than elaboration of the

structure. A vigorous dissent by Judge Dyk noted that one of ordinary skill in the art might well have found the structure of the claimed sequence adequately described and castigated the majority for ignoring the alleged infringer's failure to make any factual showing on this issue. *Id.* at 1024 (Dyk, J., dissenting). Ultimately, on the narrow issue of whether function could ever reveal structure, Judge Lourie recanted. He agreed to remand the issue to the district court for further factual findings. See *Enzo*, 296 F.3d at 1329.

138. See, e.g., Mark P. Gergen, *The Jury's Role in Deciding Normative Issues in the American Common Law*, 68 *Fordham L. Rev.* 407, 410 (1999) (noting that in negligence law, "the jury decides reasonableness . . . even when no fact is in doubt," and attributing this power allocation to the value placed on popular judgment in this area); Monaghan, *supra* note 22, at 234 ("[Q]uite plainly, the actual distribution of authority between judges and other decisionmakers has often been governed by other factors, such as the nature of the substantive issue and the character of the decisionmakers."). I discuss Gergen's analysis of the role of the jury further *infra* text accompanying notes 274–275.

law in a manner useful for future cases—should rest with the PTO and the trial courts.

A. Judicial Deference to PTO Fact Finding

1. *General Considerations.* — The normative case for deferential judicial review of fact finding by administrative agencies is grounded in considerations of expertise and procedural efficiency. As the New Deal reformers who created the administrative state recognized, courts generally lack the expertise and resources necessary to address factually complex social problems.¹³⁹ To be sure, the New Dealers' faith in the virtues of a centralized knowledge base has since given way to enthusiasm for the decentralized knowledge embodied in market transactions.¹⁴⁰ Nonetheless, as between the agencies and the courts, which are both government institutions, it is difficult to dispute that agencies possess superior resources. In addition, to the extent that an agency can get the facts "right," we can achieve a nationally uniform decision and are spared the time, expense, and delay of court proceedings. As discussed further below, the court system is an exceptionally expensive mechanism for mediating disputes, particularly disputes that turn on complicated technology.

The major argument favoring courts over agencies involves courts' ability to act in an even-handed manner that promotes overall social efficiency. A huge volume of literature, predating, but also influenced heavily by, the interest group prong of public choice theory,¹⁴¹ emphasizes

139. See, e.g., Thomas W. Merrill, *Judicial Deference to Executive Precedent*, 101 *Yale L.J.* 969, 972–75 (1992) (“[C]ourts are generalists, whereas agencies are specialists. Specialists usually have a better grasp of technical terms or the practical consequences of a decision, and thus their views should be given deference by the generalists.”); Cass R. Sunstein, *Law and Administration After Chevron*, 90 *Colum. L. Rev.* 2071, 2079 (1990) (“For the twentieth century reformers, courts lacked the flexibility, powers of coordination, initiative, democratic accountability, and expertise necessary to deal with complex social problems.”). More generally, those who created the New Deal, and sought to preserve its legacy, thought of administrative agencies as institutions that were uniquely situated to apply reason in solving complex social problems. See generally James M. Landis, *The Administrative Process* 6–46 (1938) (tracing the development of the administrative state in the United States).

140. See generally Joseph D. Kearney & Thomas W. Merrill, *The Great Transformation of Regulated Industries Law*, 98 *Colum. L. Rev.* 1323 (1998) (examining shift in goals of regulation of railroad, airline, trucking, telecommunications, electricity, and natural gas industries to promotion of competition and maximization of consumer choice).

141. Public choice theory has two related components: social choice theory and interest group theory. Social choice theory, which takes as its foundation Kenneth Arrow's classic text on voting, argues that political institutions are unlikely to reach decisions that are both coherent and representative. Kenneth J. Arrow, *Social Choice and Individual Values* (2d ed. Cowles Found. for Research in Econ. at Yale Univ. 1963) (1951). In order to achieve coherent results in situations where multiple voters have different preferences, agenda setting is crucial. But agenda setting may not reflect the underlying preferences of the voters. See Jerry Mashaw, *Greed, Chaos, and Governance* 13–14 (1997). Interest group theory is discussed further *infra* text accompanying notes 142–147.

the systematic likelihood that agencies will be influenced to take actions that are favorable to the agenda of the particular entities they regulate but unfavorable to the public interest. Interest group theory argues that, contrary to traditional accounts of the political process, the actions of political institutions are not the balanced product of lobbying by groups on all sides of an issue.¹⁴² Rather, because of the high costs associated with collective action in the form of political participation, the groups from which political actors are most likely to hear are interest groups formed around narrow agendas that are highly salient to the members of the group. In the context of regulatory action, the most vocal interest groups will be those narrow constituencies directly and immediately affected by the regulatory action.¹⁴³

The most orthodox interest group theorists view regulation as a commodity that is bought and sold to those who are willing to pay the highest price. Thus the regulated industry may offer to regulators various benefits, including the prospect of future employment. Alternatively, an interest group with influence over legislators (for example, a group that was willing to mobilize voters or make campaign contributions) may even have pressured Congress to establish the agency in the first place. The interest group would then view the agency as a mechanism for extracting rents from the public fisc.¹⁴⁴ To be sure, some of the underlying assumptions of this version of interest group theory, in particular its assumption that political actors always act to maximize their own narrow self-interest, whether in reelection resources in the case of legislators or in direct monetary gain in the case of administrators,¹⁴⁵ have been ably criticized; ideology, as well as a desire to advance the public interest, no doubt influences the actions of many political actors.¹⁴⁶ Nonetheless, one does not

142. See, e.g., Robert A. Dahl, *A Preface to Democratic Theory* 3 (1956) (“[D]emocratic theory is concerned with processes by which ordinary citizens exert a relatively high degree of control over leaders . . .”).

143. For important statements of this argument, see, e.g., Gordon Tullock, *Efficient Rent Seeking*, in *Toward a Theory of the Rent-Seeking Society* 97, 97–99 (James M. Buchanan et al. eds., 1980); Sam Peltzman, *Toward a More General Theory of Regulation*, 19 *J.L. & Econ.* 211, 212–13 (1976); George J. Stigler, *The Theory of Economic Regulation*, 2 *Bell J. Econ. & Mgmt. Sci.* 3, 10–13 (1971).

144. See, e.g., Edward L. Rubin, *Public Choice, Phenomenology, and the Meaning of the Modern State: Keep the Bathwater, but Throw Out That Baby*, 87 *Cornell L. Rev.* 309, 343 (2002) (noting that public choice theory would suggest that the “modern administrative state is created by interest group efforts to extract benefits from the state”).

145. As noted in the text, administrators may view the regulated industry as a source of future employment. Administrators may also be interested in seeking monetary gain by securing larger budgets for their agencies. See William A. Niskanen, Jr., *Bureaucracy and Representative Government* 114 (1971) (“[T]he coterminous relation of a bureaucrat’s rewards and his position implies that bureaucrat will maximize the total budget of his bureau . . .”).

146. See, e.g., Daniel A. Farber & Philip P. Frickey, *Law and Public Choice* 24–33 (1991) (arguing that “political commitment” is a valid and significant motivation for political action); Edward L. Rubin, *Beyond Public Choice: Comprehensive Rationality in the Writing and Reading of Statutes*, 66 *N.Y.U. L. Rev.* 1, 31–35 (1991) (collecting and

have to believe that political actors are concerned with only their narrow self-interest to believe that they may be influenced to some extent by the prospect of reelection resources or future employment. Indeed, even if one rejects altogether the motivational suppositions of interest group theory, it is difficult to argue with the proposition that all institutional actors—courts included—will be responsive to information provided by advocates in rough proportion to the expenditures those advocates make.¹⁴⁷ However, given the intimate interaction of many agencies with the industries they regulate, agencies may be subject to more sustained, and unchallenged, industry lobbying efforts than generalist courts.

In the abstract, it is quite difficult to ascertain the weight one should give to these arguments for and against judicial deference to agency fact finding. Fortunately, the arguments become much more clear in the context of concrete institutions. The next section addresses how these general arguments for and against judicial deference to agency fact finding apply to the relationship between the PTO and the Federal Circuit.

2. *Should the Federal Circuit Defer?*

a. *The Role of Expertise.* — Several expertise-related arguments support the Federal Circuit granting deference to PTO fact finding, particularly in the context of an appeal from a PTO patent denial. The amount of technical knowledge that an appellate court can wield—even a specialized court like the Federal Circuit—is quite limited. To be sure, the Federal Circuit has a few judges who are technically trained.¹⁴⁸ Federal Circuit judges are also assisted by a small technical staff and by law clerks who generally have both legal training and some technical background. Nonetheless, technically adept judges, a small technical staff, and technically adept clerks are not—and indeed could not be—trained in every area of science or technology in which patent disputes might arise. More-

discussing literature on legislators' varied motivations); see also Jason Scott Johnston, A Game Theoretic Analysis of Alternative Institutions for Regulatory Cost-Benefit Analysis, 150 U. Pa. L. Rev. 1343, 1361 (2002) (discussing positive political theory, a successor to public choice theory, and noting that "[a] fundamental insight of recent work in [the area] is that people who work for regulatory agencies have preferences about the programs and policies that those agencies implement"). For a general overview of positive political theory, which applies game theory to political decisionmaking, see Daniel B. Rodriguez, The Positive Political Dimensions of Regulatory Reform, 72 Wash. U. L.Q. 1, 42-110 (1994).

147. See, e.g., Susan Freiwald, Comparative Institutional Analysis in Cyberspace: The Case of Intermediary Liability for Defamation, 14 Harv. J.L. & Tech. 569, 611 (2001) (noting that Neil Komesar's participation-centered approach, which regards the motivation of institutional actors as largely irrelevant, emphasizes that all such actors will tend to respond to those parties that have the resources to make their case).

148. Judges Linn, Gajarsa, Newman, and Lourie have technical backgrounds. See U.S. Court of Appeals for the Federal Circuit, Judicial Biographies, at <http://www.fedcir.gov/judgbios.html> (last visited May 7, 2003) (on file with the *Columbia Law Review*).

over, because scientific knowledge is highly localized,¹⁴⁹ training in one area of science or technology simply does not transfer into other areas. By contrast, the PTO has thousands of patent examiners with training in multiple different fields of technology. In the area of biotechnology alone, for example, the PTO has over 150 Ph.D.s.¹⁵⁰

Indeed, in the area of biotechnology, the Federal Circuit's refusal to defer to PTO fact finding, even in the face of a relative lack of expertise, has yielded some questionable jurisprudence. As discussed above, the Federal Circuit has, in cases like *In re Deuel*, reversed PTO decisions that recognize methods as relevant prior art for DNA sequence claims.¹⁵¹ According to the Federal Circuit—or, more specifically, Judge Lourie, whose scientific training is in chemistry—because DNA sequences are like ordinary chemical compounds, the only relevant prior art for DNA sequence claims is other, structurally similar, DNA sequences.¹⁵² The court's classification of a DNA sequence as a conventional chemical compound ignores the reality that, given the genetic code, DNA is a compound that is uniquely rich in informational content.¹⁵³ In recent years, this informational role—which often requires reliance on methods rather than structure—has become increasingly central and has substantially reduced the time and expense associated with DNA sequencing and gene cloning.¹⁵⁴ The Federal Circuit's jurisprudence in this area is particularly problematic because it appears to elevate to the status of invariant legal

149. E.g., Jay P. Kesan, Carrots and Sticks to Create a Better Patent System, 17 Berkeley Tech. L.J. 763, 766 (2002).

150. In total, the PTO has about 3,000 examiners. USPTO Business Plan 10 (Feb. 2002), available at <http://www.uspto.gov/web/menu/fin03presidbudg1.pdf> (on file with the *Columbia Law Review*) (giving figure of 2,905 for calendar year 2000).

151. See *supra* text accompanying notes 74–79.

152. See *In re Deuel*, 51 F.3d 1552, 1557–58 (Fed. Cir. 1995) (“Normally, a *prima facie* case of obviousness is based upon structural similarity, *i.e.*, an established structural relationship between a prior art compound and the claimed compound.”). In ordinary chemistry, the conventional test for obviousness of a given chemical product is structural similarity to other chemical products. This test makes sense (or at least has made sense) because the properties of compounds are often difficult to predict, except to the extent that structurally similar compounds typically have similar properties. See, e.g., *In re Dillon*, 919 F.2d 688, 692–93 (Fed. Cir. 1990) (stating that rebuttal of *prima facie* case of obviousness based on structural similarity “can consist of a comparison of test data showing that the claimed compositions possess unexpectedly improved properties or properties that the prior art does not have”). Increasingly, however, informational techniques that do not rely on structure are beginning to invade ordinary chemistry as well. We should not be sanguine about the Federal Circuit altering its jurisprudence to incorporate these new techniques.

153. Kenneth J. Burchfiel, *Biotechnology and the Federal Circuit* 19 (1995) (discussing the manner in which DNA stores information necessary for making all proteins).

154. See, e.g., Arti K. Rai, *The Information Revolution Reaches Pharmaceuticals: Balancing Innovation Incentives, Cost, and Access in the Post-Genomics Era*, 2001 U. Ill. L. Rev. 173, 174–75 [hereinafter Rai, *Information Revolution*] (arguing that new, data-driven era of biotechnological research may lead to faster, more efficient drug development).

precedent a determination regarding relevant biotechnological techniques that is clearly time-specific.¹⁵⁵

To the limited extent that the Federal Circuit has acknowledged an informational role for DNA, and hence a role for methods as prior art, its awareness of the relevant technology has been inadequate. In *Deuel*, the court briefly noted that using methods that track the informational content of DNA to deduce a gene sequence from its amino acid sequence was difficult to accomplish successfully, even if obvious to attempt. The court also noted the black-letter patent principle that a risky or uncertain invention can be deemed nonobvious even if it is obvious to try.¹⁵⁶ The court suggested these arguments as an alternative ground for reaching its conclusion of nonobviousness. Contrary to the reasoning in *Deuel*, however, deducing a gene sequence from its amino acid sequence was not, even at the time the invention at issue in *Deuel* was being formulated, a particularly risky or uncertain step.¹⁵⁷

As a consequence of *Deuel*, many if not most novel DNA sequences are likely to be considered nonobvious, "however trivial the scientific advance that led to their identification."¹⁵⁸ This substantial lowering of the nonobviousness bar has led biotechnology entrepreneurs to file tens of thousands of patent applications on DNA sequences that they have been able to generate quickly through routine, automated sequencing methods.¹⁵⁹ If the PTO were to have granted patents—particularly broad patents—on large numbers of these relatively trivial upstream inventions,

155. In some respects, the Federal Circuit's biotechnology case law appears to have emulated some of the most criticized Supreme Court jurisprudence. In its cases involving broadcast regulation, specifically *Red Lion Broadcasting Co. v. FCC* and its progeny, the Supreme Court has relied on a factual determination regarding the physical "scarcity" of the broadcast spectrum to justify a different standard of First Amendment review for broadcast regulation. See 395 U.S. 367, 390 (1969). Many commentators have argued that, although the Court's factual determination may have been correct at the time of the original *Red Lion* case, it is now incorrect. See, e.g., Glen O. Robinson, *The Electronic First Amendment: An Essay for the New Age*, 47 *Duke L.J.* 899, 903-04 (1998) (criticizing the scarcity determination). Hence scarcity can no longer be used to justify special First Amendment treatment for broadcast regulation. In many ways, however, the Supreme Court's approach, though problematic, is more justified than what the Federal Circuit has been doing. The finding about spectrum scarcity was not, in contrast to the fact finding in *Deuel*, clearly incorrect at the time it was made. In addition, although the Supreme Court has not overruled *Red Lion*, the Court has indicated that it recognizes the criticism of the case and may be prepared to overrule it in the future. See Benjamin, *supra* note 37, at 283-84.

156. *Deuel*, 51 F.3d at 1558-59. For statements of this principle, see, e.g., *In re Dow Chem. Co.*, 837 F.2d 469, 473 (Fed. Cir. 1988); Robert P. Merges, *Uncertainty and the Standard of Patentability*, 7 *High Tech. L.J.* 1, 40-42 (1992).

157. See Rai, *Information Revolution*, *supra* note 154, at 836-37.

158. Rebecca S. Eisenberg & Robert P. Merges, *Opinion Letter as to the Patentability of Certain Inventions Associated with the Identification of Partial DNA Sequences*, 23 *AIPLA Q.J.* 1, 32 (1995).

159. See Eliot Marshall, *Patent Office Faces 90-Year Backlog*, 272 *Science* 643, 643 (1996) (noting that the biotechnology company Incyte claims to have filed applications on over 400,000 sequences).

there is reason to fear that downstream research might have been delayed or perhaps even blocked. The transaction costs that downstream researchers might have faced in negotiating with various upstream claimants would likely have been prohibitive.¹⁶⁰

Some commentators have argued that the *Deuel* case, as well as other cases that have significantly lowered the standard for nonobviousness in the area of biotechnology, reflect not so much the Federal Circuit's disregard for the factual findings of the PTO as the court's special solicitude for the biotechnology industry. On this view, the risk and expense of biotechnological innovation may militate in favor of a lower nonobviousness standard in that industry.¹⁶¹ The argument that the Federal Circuit should be sensitive to the economic structure of innovation in particular

160. See Michael A. Heller & Rebecca S. Eisenberg, Can Patents Deter Innovation? The Anticommons in Biomedical Research, 280 Science 698, 698–99 (1998); Arti Kaur Rai, Regulating Scientific Research: Intellectual Property Rights and the Norms of Science, 94 Nw. U. L. Rev. 77, 125–29 (1999). The argument that, even with broad upstream patents, a downstream researcher could still be required to negotiate with multiple upstream patent holders might seem counterintuitive. The reason broad upstream patents aggravate rather than alleviate transaction cost problems is that patent law allows blocking patents. Suppose, for example, that inventor A isolates a gene fragment and then receives a relatively broad patent that covers not only the fragment but also the gene of which it was a part. Inventor B might then receive a blocking patent on the gene itself. Inventor C, who wants to do research using the gene, would then have to seek a license from both A and B. Arti K. Rai, Fostering Cumulative Innovation in the Biopharmaceutical Industry: The Role of Patents and Antitrust, 16 Berkeley Tech. L.J. 813, 837–38 (2001) [hereinafter Rai, Cumulative Innovation]. In contrast, if inventor A's patent had been limited to the gene fragment itself, inventor C, who wanted to use the gene, would probably have to negotiate with inventor B only.

Fortunately for the progress of biomedical research, the Federal Circuit's subsequent decisions limiting patent scope for DNA patents, as well as more general guidelines on patentability recently issued by the PTO, make it unlikely that the PTO will in fact grant broad patents on gene sequences of unknown function. See Utility Examination Guidelines, 66 Fed. Reg. 1092, 1097–99 (Jan. 5, 2001); Guidelines for the Examination of Patent Applications Under the 35 U.S.C. 112, ¶ 1, "Written Description" Requirement, 66 Fed. Reg. 1092, 1104–07 (Jan. 5, 2001). The utility guidelines, which raise the level of functionality the patentee must show for her invention, primarily respond to arguments about excessive patenting made by the community of academic scientists. The written description guidelines, which narrow patent scope, are primarily a response to the Federal Circuit decision in *Regents of the University of California v. Eli Lilly & Co.*, 119 F.3d 1559 (Fed. Cir. 1997), discussed *supra* text accompanying notes 136–137. Thus the Federal Circuit has, to some extent, unwittingly mitigated the very problem it created. However, as discussed further below, this "mitigation" may have problematic consequences of its own. See *infra* text accompanying notes 163–167.

161. See, e.g., Karen I. Boyd, Nonobviousness and the Biotechnology Industry: A Proposal for a Doctrine of Economic Nonobviousness, 12 Berkeley Tech. L.J. 311, 337–39 (1997) (proposing that an invention in the biotechnology field "need only be technically *or* economically nonobvious to be patentable"); Robert P. Merges, One Hundred Years of Solicitude: Intellectual Property Law, 1900–2000, 88 Cal. L. Rev. 2187, 2225–26 (2000) (noting that "[w]ith respect to nonobviousness, observers widely agree that the Federal Circuit has been very generous toward [biotechnology industry] patentees and patent applicants").

industries has considerable merit.¹⁶² The suggestion that cases like *Deuel* reflect an economic rationale is, however, puzzling. In the first instance, the Federal Circuit's opinions offer no hint of such a rationale. In addition, to the extent that the court is adopting an economic rationale *sub silentio*, it would have to be invoking not the conventional argument that patents of some sort provide the necessary incentive to *find* gene sequences (finding gene sequences is an automated and inexpensive process, and has been for some time) but, rather, the more controversial claim that *broad* patents on early-stage inventions like gene sequences provide the necessary economic incentive for *developing* those sequences into marketable products.¹⁶³ Even assuming that the latter rationale has merit,¹⁶⁴ Federal Circuit decisions like *Regents of the University of California v. Eli Lilly & Co.*, which apply patent disclosure requirements stringently so as to mandate that gene patents be quite *narrow* in scope, strongly indicate that the court is not following that rationale.¹⁶⁵ Indeed, the Federal Circuit has, in the area of biotechnology, produced a jurisprudence that is at odds with the economic structure of the industry. Under the court's case law, patents are easily secured, but they are very narrow in scope, regardless of the stage of research that they cover. Very narrow patent scope is not a problem for patents on research that is quite upstream. Indeed, if we are going to allow patents on such upstream, fundamental research—as the Federal Circuit has apparently decided to do—narrow scope may be the only mechanism available for avoiding transac-

162. See *infra* Part III (proposing broader policy role for the Federal Circuit).

163. The argument that broad, early-stage patents are useful for motivating development has been put forward by some patent scholars, perhaps most notably Edmund Kitch. See Kitch, *supra* note 10, at 267–71, 276 (arguing that absent broad patent protection early in the development process, no one will invest for fear that “the fruits of the investment will produce unpatentable information appropriable by competitors”).

164. In keeping with the Article's institutional focus, I do not reach a conclusion on the ultimate merits of this rationale. The rationale does have some force, particularly in the area of biotechnology, which is notorious for its high costs of development. On the other hand, broad patents may hinder innovation to the extent that the owner of the broad patent may encounter transaction cost difficulties in licensing the patent widely, to a variety of follow-on improvers. See Robert P. Merges & Richard R. Nelson, On the Complex Economics of Patent Scope, 90 Colum. L. Rev. 839, 885–91 (1990). These transaction cost difficulties may also arise to the extent that multiple parties have overlapping claims in a basic research platform. See *supra* note 160. Further, broad patents in early-stage invention may be unnecessary to the extent that researchers are likely to be able to secure additional patents as research moves down the R&D path. For example, in the context of biopharmaceuticals, patents on the “new chemical entity” that represents the active ingredient in a drug are uncontroversial.

165. See 119 F.3d at 1566–69. Notably, the *Lilly* case rests squarely upon the Federal Circuit's erroneous fact finding in cases like *Deuel*. To the extent that a method for finding a DNA sequence cannot render the sequence obvious, merely citing such a method cannot allow one to claim the sequence. Thus biotechnology patentees will be quite limited in what they can claim. See Rai, Addressing New Technology, *supra* note 135, at 835–37. The *Lilly* case thus illustrates the manner in which a low standard for nonobviousness can lead to a high standard for disclosure.

tion cost bottlenecks in follow-on research.¹⁶⁶ But narrow scope does become a problem as research moves further downstream. For example, if patent claims on compounds that must be taken through the FDA's drug approval process are interpreted very narrowly, they will not give firms—particularly small firms—the monopoly power they need to provide a hedge against the cost and risk of this process.¹⁶⁷

From an economic standpoint, a more appropriate approach might involve an explicit recognition on the part of the court that patents should be allowed a wider scope as research moves further downstream. Indeed, as discussed further in Part III, an economic argument for increasing patent scope as research moves further downstream can be made for all cumulative research. The lack of fit between the Federal Circuit's jurisprudence and biotechnology's economic structure (or, indeed, the economic structure of cumulative innovation more generally) is not surprising. After all, the court's opinions as a whole appear to betray little interest in economic or policy analysis. In all likelihood, then, the Federal Circuit's refusal to defer to PTO fact finding does not reflect an economic rationale. In any event, even assuming *arguendo* that the Federal Circuit's biotechnology jurisprudence has some economic motivation, its failure to enunciate this motivation precludes any meaningful development of patent policy by the court. As discussed further in Part III, judicial policy development requires that economic arguments be considered explicitly. Without such explicit consideration, there can be no policy evaluation and evolution.

Other defenders of the Federal Circuit have suggested that its biotechnology jurisprudence may represent a counterweight to the tendency of scientists and other decisionmakers to view invention as obvious in hindsight, particularly if the obviousness determination is reached many years after the invention was actually made.¹⁶⁸ Indeed, hindsight bias is a

166. See Rai, *Cumulative Innovation*, *supra* note 160, at 853.

167. For example, the degeneracy of the genetic code creates a situation where a particular therapeutic protein may be coded for by many different DNA sequences. If a patentee claims the protein as a DNA sequence, but is allowed to claim only those DNA sequences that it has actually isolated, a rival may be able to "invent around" the therapeutic protein patent by making trivial changes to the claimed DNA sequences. If such inventing around is successful, it will create competition that severely undermines the patentee's return on its R&D investment. For empirical studies showing that patents are particularly critical for appropriating returns to R&D in the pharmaceutical arena, see Wesley M. Cohen et al., *Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not)* 3 (Nat'l Bureau of Econ. Research, Working Paper No. 7552, 2000), available at <http://www.nber.org/papers/w7552> (on file with the *Columbia Law Review*); Richard C. Levin et al., *Appropriating the Returns from Industrial Research and Development*, 3 *Brookings Paper on Economic Activity* 783 (1987).

168. See, e.g., Lawrence M. Sung, *Stranger in a Strange Land: Biotechnology and the Federal Circuit*, 2 *Wash. U. J.L. & Pol'y* 167, 170–72, 176–77 (2000) (discussing the Federal Circuit's obviousness jurisprudence in the area of biotechnology in light of the significant

problem that all decisionmakers must confront,¹⁶⁹ and it is one against which the Federal Circuit has inveighed.¹⁷⁰ Patent law requires that non-obviousness be judged as of the time of invention. In cases like *In re Deuel*, however, there is no reason to believe that PTO examiners were guilty of hindsight bias. To the contrary, as noted earlier, the identification of coding gene sequences through specific probing techniques had become routine by the time that the invention at issue in *Deuel* was made in the mid-1980s.¹⁷¹

Finally, one might argue that although the Federal Circuit's decisions in the area of biotechnology (and, as discussed above, computer software¹⁷²) are divorced from technological reality, they have the virtue of predictability. The decisions set up bright-line rules, and parties can arrange their conduct around such rules. This argument has considerable explanatory force. Indeed, given the court's general embrace of bright-line rules,¹⁷³ the argument probably describes the motivation of many Federal Circuit judges. As a normative matter, however, the substitution of bright-line rules for fact-specific standards is unwise. Such substitution assumes that initial allocations of patent rights do not matter, so long as they are predictable. To be sure, predictability is a very important value in patent law—but it is not the only value. To the contrary, all the major economic theories of patent law assume that the various incentive effects created by initial rights assignment (and/or transaction costs associated with reassignment) are sufficiently important that it matters how we define and assign such rights.¹⁷⁴ In addition, although bright-line

delay between the date of filing of a patent application and the resolution of an infringement action).

169. See, e.g., Christine Jolls et al., A Behavioral Approach to Law and Economics, 50 *Stan. L. Rev.* 1471, 1523 (1998) (discussing danger of hindsight bias in jury determinations of negligence); Jeffrey J. Rachlinski, A Positive Psychological Theory of Judging in Hindsight, 65 *U. Chi. L. Rev.* 571, 571–72, 613–15 (1998) (defining hindsight bias as the tendency to assign an erroneously high probability to a particular event simply because it actually occurred, and discussing the use of secondary considerations as a mechanism for guarding against hindsight bias in the context of nonobviousness determinations).

170. See, e.g., *In re Deuel*, 51 F.3d 1552, 1558 (Fed. Cir. 1995) (criticizing the PTO for “impermissible hindsight reconstruction of the claimed invention”).

171. Arguably, these techniques had become relatively routine even earlier. In *Fiers v. Revel*, an interference proceeding involving the question of who had priority to a DNA sequence encoding fibroblast beta-interferon (a protein that promotes viral resistance in human tissue), one of the parties argued that, as of 1979, identifying and cloning a gene by using a nondegenerate probe to screen a small DNA library had become “routine to those skilled in the art.” 984 F.2d 1164, 1168 n.9 (Fed. Cir. 1993).

172. *Supra* text accompanying notes 128–133.

173. See *infra* Part III.A.

174. Scholars such as Edmund Kitch, who regard patent rights as similar to ordinary property rights, argue that broad rights should be given early in the R&D process. See Kitch, *supra* note 10, at 267–71, 276. On this view, broad rights will give the innovator an incentive to develop the invention in an efficient manner. Moreover, although Kitch does not expressly address the question, he appears to assume that the transaction costs of accumulating narrow rights into a broad right are such that the initial allocation should be

rules may increase predictability for future cases, they can upset expectations quite dramatically when they are first enunciated.¹⁷⁵

A policy of deference to PTO fact finding in the context of patent denials may also discipline any institutional tendency that the Federal Circuit has towards pro-patent bias. As discussed further in Part III, the Federal Circuit, unlike generalist courts, hears disproportionately from members of the patent bar. Although patent bar members represent both patent plaintiffs and patent defendants, both plaintiffs and defendants are often patent holders.¹⁷⁶ Thus, while a patent lawyer may attack the particular patents held by her opponent, she is unlikely to make more sweeping legal and policy arguments that emphasize the problems caused by strong, or numerous, patents. Indeed, the organized intellectual property bar has tended to decry the recent questioning of whether patents have become too powerful.¹⁷⁷

By the same token, any enthusiasm about the PTO's fact-finding expertise must be confined to the context of patent denials. In contrast, when patents are *granted*, there is reason to believe that various institutional pressures in favor of patent grants may have outweighed the advantages of PTO expertise. As discussed in the next section, these pressures extend beyond the usual concerns identified by interest group theorists.

b. *Patent Grants: The Role of Pro-Patent Bias.* — As an initial matter, it bears emphasis that examiners are compensated in part based on the number of final dispositions of patents that they accumulate.¹⁷⁸ Because it is easier and faster to secure a final disposition by allowing a patent application than by denying it, there is an incentive to allow applications.¹⁷⁹ The incentive to grant rather than reject patents is exacerbated

broad. Scholars who follow the conventional view that patents serve not as ordinary property but primarily as incentives to invent argue that rights should be given somewhat later in the R&D process and should not be overly broad. For further discussion of these theories, see *infra* notes 353–355 and accompanying text.

175. For example, *Regents of the University of California v. Eli Lilly & Co.*, 119 F.3d 1559 (1997) (discussed *supra* note 165 and accompanying text), upset expectations in the biotechnology community quite dramatically when it was decided. See Janice M. Mueller, *The Evolving Application of the Written Description Requirement to Biotechnological Inventions*, 13 *Berkeley Tech. L.J.* 615, 652 (1998) (noting that “the *Lilly* decision . . . reflect[s] an increasingly-widening gulf between the norms of the business and scientific community and those of the United States patent system”).

176. See *infra* note 209 (discussing widespread patenting in most industries, often for purely defensive reasons).

177. See *infra* note 340 (discussing statement of American Bar Association Intellectual Property Committee Chair).

178. *Merges*, *Six Impossible Patents*, *supra* note 5, at 607 (stating that examiner bonuses are calculated based upon the number of final dispositions).

179. *Id.* As John Thomas explains,

Patent Office management tracks examiner workloads by crediting either a first action on the merits or a disposal of the application. Disposal counts are generally awarded either for the allowance or abandonment of the application. Noticeably absent from the list of ways to obtain a disposal count is continued rejection of the application.

by the fact that the examiner bears the burden of proof with respect to invalidity and must write up reasons for rejection but not reasons for allowance.¹⁸⁰ More generally, because the patent office is entirely supported by fees paid by patent applicants,¹⁸¹ it has an institutional tendency to want to treat those applicants well. Perhaps above all, because the administrative process for securing a patent is largely conducted as an *ex parte* proceeding,¹⁸² the PTO is even more subject than usual to unchallenged lobbying by the entity interested in securing the regulatory dispensation. The lack of a proceeding is a particular problem in the context of software patents, including software-implemented business method patents, because the prior art is not organized in a manner that is readily accessible to examiners.¹⁸³ The absence of an administrative niche for challenging parties persists despite the fact that third parties who compete with the patentee are most likely to have the information and interest necessary to challenge a patent.¹⁸⁴ The net result of these problems is that, when patents are finally challenged in the adversarial context of a judicial proceeding, a large percentage are invalidated based on prior art that was not cited before the PTO.¹⁸⁵

Thomas, *supra* note 5, at 324.

180. See *id.* at 325.

181. John F. Duffy, *Harmony and Diversity in Global Patent Law*, 17 *Berkeley Tech. L.J.* 685, 700 (2002).

182. Since 1980, a reexamination procedure for issued patents has been in existence. See Act of Dec. 12, 1980, Pub. L. No. 96-517, 94 Stat. 3015 (codified at 35 U.S.C. §§ 301-307 (2000)). Under this procedure, a private party that is unhappy with a particular patent grant can request that the PTO conduct a reexamination. These reexaminations have traditionally been conducted in an *ex parte* fashion, however. Moreover, although Congress recently authorized the creation of an *inter partes* reexamination procedure, this procedure is limited by the fact that it precludes the third-party challenger from attacking the patent subsequently in a judicial setting. See Mark D. Janis, *Inter Partes Patent Reexamination*, 10 *Fordham Intell. Prop. Media & Ent. L.J.* 481, 492-98 (2000). In addition, until recently, only the patentee could appeal the results of the *inter partes* reexamination. See *id.*

183. While prior patent applications and academic journal publications are the type of prior art that can most readily be searched, much of the relevant prior art in the software and business methods area takes the form of informal publications in trade journals. Of course, the difficulty of searching for prior art in the area of business method patents would have been substantially mitigated had the Federal Circuit displayed more foresight in defining the circumstances under which such patents could be granted. See *infra* text accompanying notes 305-323.

184. Merges, *Six Impossible Patents*, *supra* note 5, at 615.

185. See Allison & Lemley, *supra* note 56, at 234 (observing that in cases of prior art-based invalidation, courts relied on prior art not considered by the PTO approximately fifty-eight percent of the time, with the result that "reliance on uncited art was more likely to lead to a finding of invalidity than reliance on cited art"). A proposal for PTO reform put forward by Jay Kesan and Mark Banik argues that information disclosure to the PTO can and should be stimulated by requiring that courts refrain from invalidating patents based on disclosed prior art unless they are convinced that no reasonable examiner would have allowed the patent in light of the prior art. Kesan & Banik, *supra* note 5, at 26 & n.10. However, the data collected by Allison and Lemley indicate that, when no new prior art is cited, courts already give patents a fairly strong presumption of validity. Allison & Lemley,

These structural problems give us reason to be wary of deference in the context of patent grants. As a consequence, the Federal Circuit's longstanding policy of deference to patent grants but not to patent denials appears to have gotten it precisely backwards. In order to merit at least some deference in the context of patent grants, the PTO's system for fact finding needs to be revised. Although the PTO should not necessarily expend large sums in examining every patent application, it should be endowed with greater expertise. A straightforward, and relatively inexpensive, mechanism for focusing such expertise on the most commercially significant patents would be the adoption of an adversarial opposition proceeding. As various commentators have pointed out, this type of opposition proceeding would remedy the information deficit faced by the PTO by bringing into the system private knowledge about the prior art.¹⁸⁶ To the extent that the patent interferes with their own product development plans, competitors of a given patent applicant should have a significant monetary incentive to bring forward information that undermines the patent. Moreover, if a low cost administrative proceeding can defeat the patent, it is preferable to the alternative of challenging validity in a high cost judicial infringement action.¹⁸⁷

supra note 56, at 234. Giving an even greater preference for cited prior art would encourage patentees to overwhelm the patent office with prior art. Under the Kesan and Banik proposal, all such art would weigh in favor of the patentee, irrespective of whether the patent examiner had any opportunity to evaluate it carefully. See Kesan, supra note 149, at 763–64.

186. See, e.g., Kesan, supra note 149, at 776; Merges, Six Impossible Patents, supra note 5, at 615. One relatively minor point of dispute among those who support an opposition proceeding is whether the proceeding should be held before or after the patent is granted. For purposes of my argument, precisely when the opposition proceeding is held is not critical.

187. See Jonathan Levin & Richard Levin, Patent Oppositions 5 (John M. Olin Program in Law and Economics Working Paper No. 245, SIEPR Discussion Paper No. 01-29, 2002), available at http://papers.ssrn.com/paper.taf?abstract_id=351900 (on file with the *Columbia Law Review*) (arguing that “the costs of introducing an opposition system are likely to be small in relation to the potential benefits”). The authors note that even in a model that includes only a single inventor, a single potential infringer, and their respective customers (and hence does not include positive externalities accruing to other parties from greater certainty and more timely information about the likely validity of patents), introducing an opposition process will have an unambiguous welfare benefit in those cases where the process obviates the need for litigation. *Id.* at 4–5. In those cases where litigation would not have occurred in any event because the potential infringer “might either have refrained from development or been able to negotiate a license without litigation,” oppositions could create deadweight loss. *Id.* at 4. However, this static deadweight loss would be counterbalanced by static benefits in the form of deadweight loss avoided by consumers and by dynamic benefits produced because rewards are accurately tailored to “the true novelty and nonobviousness of the [] invention.” *Id.*

John Thomas argues that competitors contemplating an opposition proceeding might face a collective action problem. Although competitors of the patentee will have an interest in challenging patents that interfere with their own product development plans, no single competitor may want to expend the money necessary to challenge a patent, or patent application, given that other competitors will be able to free-ride off the single competitor's efforts. See Thomas, supra note 5, at 333–40. Thomas argues that some type

In addition, even though competitors of a given patentee who are themselves patent holders might be unwilling to make legal or policy arguments lowering the overall level of patent protection in a given industry, they should have no qualms about making the factual arguments necessary to defeat a *particular* patent that poses an obstacle to their business plans. To put the point another way, when there are arguments that might undermine a particular patent without weakening the patent system more generally, an opposition proceeding is likely to bring these arguments to the attention of the PTO. As a consequence, even though the PTO, like many agencies, may tend not to hear from those who are opposed to the grant of regulatory favors,¹⁸⁸ opposition proceedings will provide some opportunity for the airing of contrary views.¹⁸⁹ Notably, opposition proceedings have been used successfully for many years in the European Patent Office and in the patent system of the United Kingdom.¹⁹⁰ Opposition proceedings are the single most important vehicle for bringing the traditional administrative law virtues of neutrality and expertise to bear upon patent grants.¹⁹¹

of additional bounty incentive should be given to the competitor that successfully opposes a patent. *Id.* at 340–42. Presumably, however, we should first institute opposition proceedings before we determine whether a bounty is necessary. Additionally, it is not clear why collective action difficulties should pose a bigger problem in the context of opposition proceedings than in the context of potential infringement litigation. In the latter context, any given infringer should be willing to pay a license fee rather than to shoulder single-handedly the financial burden of challenging a patent's validity. So if, as Thomas suggests, some type of bounty incentive is necessary in the context of opposition proceedings, such an incentive would presumably also be necessary in ordinary infringement proceedings.

188. See *supra* notes 142–144 and accompanying text.

189. Relatedly, because opposition proceedings will focus on questions of fact in individual patent cases, the concept of implementing such a reform should not be so threatening to patent holders that they would obstruct implementation in Congress. As discussed further below, because Congress tends to be influenced disproportionately by patent holders, certain reforms, such as abolition of the Federal Circuit or wholesale restructuring of the PTO, may not be politically feasible. See *infra* Part III.D.1.

190. See Eur. Patent Off., *How to Get a European Patent* 29 (9th ed. 1992); Patents Act, 1977, c. 37, § 72(2)(b) (Eng.). A recent study of European Patent Office (EPO) administrative opposition proceedings, which are allowed for nine months after a patent has issued, found that 8.3% of EPO patents are challenged in opposition proceedings; by contrast, only 0.3% of PTO patents are reexamined. Stuart J.H. Graham et al., *Patent Quality-Control: A Comparison of U.S. Patent Reexaminations and European Patent Oppositions* 14 (NBER Working Paper No. 8807, Aug. 5 2002, revised Nov. 2, 2002), available at http://www.emlab.berkeley.edu/users/bhhall/GHHM_Nov02.pdf. Over 35% of challenged EPO patents are revoked, and many of the remainder have their claims amended. *Id.* at 38 tbl.5. The study also found that opposition proceedings are less costly than litigation. *Id.* at 8–10.

191. Indeed, in its recent opinions discussing deference to agencies, the Supreme Court has also indicated that public, trial-type procedures are particularly likely to produce the sort of agency decisionmaking to which judges should defer. See, e.g., *United States v. Mead Corp.*, 533 U.S. 218, 230 (2001) (“It is fair to assume generally that Congress contemplates administrative action with the effect of law when it provides for a relatively formal administrative procedure tending to foster the fairness and deliberation that should

Moreover, to alleviate institutional tendencies toward granting patents, the financial structure of the PTO should also be changed in several ways. First, the PTO should be supported by general revenue, not by the fees paid by patent applicants, a significant percentage of which are currently siphoned off by Congress in any event.¹⁹² Second, examiners who deny grants should be rewarded to the same extent as those who grant patents.

The fact finding of PTO examiners with improved capabilities and more financial independence should merit some deference, even in the context of patent grants. Deference should, however, be given on a sliding scale. The appropriate level of deference should be high if the patent has gone through an opposition proceeding. In that context, the Federal Circuit's current position that the grant be overturned only by "clear and convincing evidence"¹⁹³ might be merited, at least for purposes of factual arguments that were actually before the PTO. As for those patents that had not gone through opposition proceedings, a significantly lower level of deference would be appropriate. Even with respect to these patents, however, a low-level presumption of validity—for example, placing the burden of proof on the challenger with respect to factual evidence actually presented to the examiner—might be merited: Competitors may not have instituted an opposition proceeding because the quality of the patent was high from the outset.

c. *The Virtues of (Relatively) Early Certainty.* — Greater deference to fact finding by a newly formulated PTO would yield benefits beyond accuracy in decisionmaking. It would also allow the relevant rights to be defined from the outset with a greater degree of certainty. To the extent patent rights are clear from the outset, the patent system can serve its intended function of providing an incentive to create and develop invention. By contrast, uncertain patent rights are of diminished value in mitigating the risk associated with innovation. Certain patent rights are important not only for patentees but also for competitors. Well-delineated patent rights provide competitors with low cost information about the strength of the patentee's position in a given area of innovation. By contrast, uncertain patent rights either chill legitimate inventive activity or force competitors to engage in costly information gathering and/or litigation to assess the validity of the patent right.¹⁹⁴

underlie a pronouncement of such force."). Although the Court's analysis has focused on the weight judges should give to agency legal interpretations, the policy argument presumably applies equally, if not more strongly, to findings of fact.

192. See U.S. Patent and Trademark Off., *The 21st Century Strategic Plan 3* (Feb. 3, 2003), available at http://www.uspto.gov/web/offices/com/strat21/stratplan_03feb2003.pdf (on file with the *Columbia Law Review*) (noting concern about diversion of fees).

193. See *supra* note 92 and accompanying text.

194. Cf. Kimberly A. Moore, *Forum Shopping in Patent Cases: Does Geographic Choice Affect Innovation?*, 79 N.C. L. Rev. 889, 927–31 (2001) [hereinafter Moore, *Forum Shopping*] (discussing possibility that uncertainty improperly decreases or increases the value of the patent right); Craig Allen Nard, *Certainty, Fence Building, and the Useful*

Emphasizing the virtues of early confidence in patent rights may appear to do little more than confirm what is already conventional wisdom. In recent years, however, some commentators have questioned the advantages of such certainty. Mark Lemley has argued that because many patents do not serve the familiar goals of providing incentives for creation and development (and indeed, sometimes do not have much relevance for industry at all), examination of patents *ex ante* may ultimately be less cost effective than litigation over, or licensing of, a few important patents *ex post*.¹⁹⁵ Ian Ayres and Paul Klemperer have questioned the value of certainty by suggesting that uncertainty may reduce the deadweight loss associated with patent protection.¹⁹⁶ I address each of these arguments in turn.

Mark Lemley invokes a rough cost-benefit analysis to argue against reform that would give the PTO a central role. Lemley argues that the costs associated with improving the PTO examination process would exceed any benefits that could be realized by eliminating “bad” (that is, poor quality) patents.¹⁹⁷ The key to Lemley’s argument is that bad patents do not pose a significant problem because, like patents generally, most bad patents are neither litigated nor licensed.¹⁹⁸ In other words, potential infringers do not even have to pay attention to most bad patents. Although Lemley does not frame his analysis in institutional terms, he essentially asserts that the combination of market institutions and the courts will be more cost effective in sorting out poor quality patent applications than will an administrative agency like the PTO.¹⁹⁹

Lemley’s argument is a significant one. If improvement in the PTO’s fact-finding capability did require additional resources that exceeded downstream benefits, such improvement would be problematic. In other words, if early administrative resolution of a particular regulatory puzzle, coupled with deference by other decisionmakers in the system, does not decrease overall expenditure, then there is reason to question the superiority of an administrative resolution.

Lemley bases his analysis on a few empirical assertions. Some of these assertions emerge from relatively hard data: Lemley notes, for example, that about two percent of patents ultimately become the subject of

Arts, 74 *Ind. L.J.* 759, 759–60 (1999) (discussing the patentee’s need for “proprietary certainty” and the competitor’s need for “competitive certainty”).

195. See Mark A. Lemley, *Rational Ignorance at the Patent Office*, 95 *Nw. U. L. Rev.* 1495, 1500–08 (2001).

196. Ian Ayres & Paul Klemperer, *Limiting Patentees’ Market Power Without Reducing Innovation Incentives: The Perverse Benefits of Uncertainty and Non-Injunctive Remedies*, 97 *Mich. L. Rev.* 985, 986–87 (1999).

197. Lemley, *supra* note 195, at 1500–08.

198. *Id.* at 1510–11.

199. *Id.* at 1497. Because Lemley does posit an important role for courts, his argument is not a simple assertion that initial allocations of rights do not matter because market players will rearrange these rights. For discussion of this argument, see *supra* text accompanying notes 173–175.

litigation, and that the cost of a full blown trial and appeal is about \$1.5 million for each side.²⁰⁰ Other data are softer: Lemley has to guess, for example, that about five to fifteen percent of patents are actually licensed for a royalty fee.²⁰¹ He also has to guess that the transaction costs of licensing are about fifty to eighty thousand dollars per license.²⁰² Based on these data, Lemley calculates the cost of litigation and licensing outside of litigation as ranging anywhere between \$2.63 billion and \$4.76 billion.²⁰³

Lemley then calculates the costs and benefits that might be associated with particular improvements in the patent examination process. Specifically, he estimates the cost of doubling the average amount of time that examiners spend on each patent, from an average of eighteen to thirty-six hours, at \$1.52 billion.²⁰⁴ He then estimates that such doubling would reduce the percentage of patents that had to be litigated or licensed by ten percent.²⁰⁵ Assuming that this ten percent decrease would affect only “bad” patents, we would save between \$262 million and \$476 million in unnecessary litigation and licensing (that is, ten percent of the \$2.63 billion to \$4.76 billion in costs that issued patents impose). Because even a cost savings, or benefit, of \$476 million is clearly less than a cost of \$1.52 billion, Lemley concludes that reform that involved substantial increases in the amount of time PTO examiners devoted to examination would not be cost effective.²⁰⁶

Lemley’s analysis is an admirable effort to provide a quantitative sense of the role that patents play. His analysis also reinforces the important point, also suggested by various empirical surveys, that in many industries patents are not considered particularly effective mechanisms for appropriating returns to R&D investment.²⁰⁷ Ultimately, however, the analysis suffers from several limitations. As an initial matter, Lemley fails to acknowledge costs unrelated to litigation and licensing that patents can impose. Most obviously, in cases where a potential user cannot afford to litigate a patent or license it at a supra-competitive price, the presence

200. Lemley, *supra* note 195, at 1501–02.

201. *Id.* at 1509–10, 1512.

202. *Id.* at 1507, 1512. Lemley believes that the actual level of licensing falls at the lower bound of this estimate. He acknowledges, however, there are no good data on licensing. Therefore, for purposes of his analysis, he varies his assumptions regarding licensing considerably, up to an estimate of fifteen percent. See *id.* at 1512.

203. *Id.* at 1509, 1512.

204. *Id.* at 1510.

205. *Id.* at 1509–10, 1512.

206. *Id.* at 1512.

207. See, e.g., Cohen et al., *supra* note 167, at 3–4 (noting that firms report patents “to be relatively ineffective (for most industries) in protecting returns to innovation”); Levin et al., *supra* note 167, at 812–19 (indicating that patents are not very important mechanisms for recouping R&D dollars in most industries, with the notable exception of the pharmaceutical and chemical industries).

of the patent may chill use.²⁰⁸ Use may also be chilled if the transaction costs of licensing are too high. Lemley's analysis does not account for the possibility of valuable uses that do not go forward because of the presence of patents.

Further, Lemley assumes that the large amount of effort devoted to defensive patenting—that is, to the accumulation of patents simply for purposes of asserting them against competitors who threaten to sue for patent infringement²⁰⁹—is costless. This assumption is questionable. Smaller firms that do not have, or cannot afford to amass and maintain, a portfolio of defensive patents may be dissuaded from entering a particular research area.²¹⁰ In addition, even for large firms, the opportunity cost of accumulating and maintaining defensive portfolios, and of assessing the credibility of competitors' portfolios, can be quite high.²¹¹ Indeed, absent evidence that patent portfolio races encourage inter-firm exchanges of information that would otherwise not have taken place, portfolio races would appear to represent a social cost. Lemley also assumes that even though patents play an important role in attracting venture capital for start-up concerns, by assuring investors that the start-ups have some level of exclusivity (or at least freedom to operate) in their chosen area of research,²¹² low-quality patents will not create bad signals

208. This standard problem with patents might be mitigated if the patentee is willing and able to engage in price discrimination. Because of difficulties associated with sorting consumers according to willingness to pay, price discrimination, however, can be difficult to implement successfully. See Hal R. Varian, *Microeconomic Analysis* 241 (3d ed. 1992).

209. Defensive patenting has become particularly prominent in certain industries like the semiconductor industry, where innovation is cumulative, and a thicket of relevant patents often exists. See Bronwyn H. Hall & Rosemarie Ham Ziedonis, *The Patent Paradox Revisited: An Empirical Study of Patenting in the U.S. Semiconductor Industry, 1979–1995*, 32 *RAND J. Econ.* 101, 104 (2001). In such industries, firms gain freedom to operate through defensive patenting. Indeed, within the semiconductor industry, it appears that much of the increase in patenting per R&D dollar over the last two decades has been the consequence of defensive patenting. See *id.* (noting that firms “appear to be engaged in ‘patent portfolio races’ aimed at reducing concerns about being held up by external patent owners and at negotiating access to external technologies on more favorable terms”). According to Hall and Ziedonis, this increase in defensive patenting “is causally related to the pro-patent shift in the U.S. legal environment in the 1980s.” *Id.* I discuss this alleged “pro-patent” shift *infra* Part III.B.

210. Hall and Ziedonis report that, in the semiconductor industry, a new manufacturer without a defensive patent portfolio might need to pay \$100 to \$200 million of revenue to license “what are now considered basic manufacturing principles.” Hall & Ziedonis, *supra* note 209, at 109–10. See also Harold C. Wegner & Stephen B. Maebius, *Patent Flooding: America's New Patent Challenge 4* (Apr. 11, 2002), at <http://www.ftc.gov/opp/intellect/020411maebius.pdf> (on file with the *Columbia Law Review*) (“[W]here a company has a huge defensive patent portfolio, while it may need that portfolio to cross-license equally large players in a field, it could also use that portfolio to exert a patent tax against newcomers and the smaller players in the field.”).

211. The opportunity cost is even higher if the defensive portfolio race requires a firm to focus on research on which it would not otherwise have focused.

212. Whether the assurance of exclusivity or freedom to operate is the only signal provided by a patent is an open question. Clarisa Long argues that patents may serve as

for investors. He argues that sophisticated venture capitalists will be able to discern with relative accuracy which patents are valid and which are not.²¹³ Even assuming that venture capitalists are adept at sorting the wheat from the chaff, such sorting has costs. In addition, to the extent that venture capitalists fail to sort accurately, they may invest in firms that fail and then sour on a particular area of technology entirely. In that situation, firms with legitimate patents may be tarred with the same brush as those with illegitimate patents.

Admittedly, it is difficult to quantify the costs imposed by the chilling effect of patents or by portfolio races and other nontraditional uses of patents. Thus, pointing out these additional costs does not provide an entirely satisfactory response to Lemley's quantitative analysis. But even if we accept *arguendo* Lemley's numbers regarding the costs imposed by patents—in other words, if we restrict ourselves to Lemley's estimate of between \$2.63 and \$4.76 billion—his analysis is questionable. Lemley assumes that doubling the number of examiner hours would result in only a ten percent reduction in the number of patents that had to be litigated or licensed.²¹⁴ This figure is dubious, particularly if we assume that all increases in examiner time were devoted to third-party opposition proceedings. Doubling examiner time through the mechanism of third-party opposition proceedings would presumably mean that examiners would be focusing on those patents that were of poor quality but nonetheless commercially significant in that they would create significant impediments to competitors' R&D plans (and hence would lead to unnecessary licensing or litigation). If increased examiner time were to be channeled primarily into opposition proceedings, the percentage of litigation and licensing might be reduced by a figure much higher than ten percent. Indeed, given that up to fifty percent of patents that are challenged in litigation are invalidated,²¹⁵ a figure of fifty percent is not entirely unrealistic. Using Lemley's own figures, a fifty percent reduction in litigation and licensing costs would translate into \$1.3 to \$2.4 billion in savings. These savings might well justify an expenditure of \$1.52 billion. More importantly, through third-party opposition proceedings, we might reduce unnecessary litigation and licensing without increasing examiner expenditures significantly. Examiners might simply channel some percentage of the time they currently spend reviewing applications *ex parte* into the adversarial context. For example, if examiner expenditures were

signals for other firm characteristics. Clarisa Long, *Patent Signals*, 69 U. Chi. L. Rev. 625, 627–28 (2002). For example, even if the patents in question do not secure rent streams, they may assure venture capitalists that a firm is conducting interesting research or is otherwise well-managed. *Id.* at 653. Long acknowledges, however, that mere patent counts are noisy signals for conveying information, even about firm characteristics unrelated to exclusivity. See *id.* at 654. Thus, even under Long's theory of signaling, investors will have to engage in costly evaluation of individual patents.

213. See Lemley, *supra* note 195, at 1514–15.

214. *Id.* at 1496 n.4.

215. *Id.*

increased by about fifty percent, or an additional \$750 million, a mere twenty-five percent reduction in a number of patents that had to be litigated or licensed would more than justify the additional expenditure.

To be sure, third-party opposition proceedings would require some expenditure on the part of private competitors of the patentee. It is likely, however, that these costs would be significantly smaller in the administrative setting than in the context of full blown litigation. A patent that takes \$1.5 million to defeat in the context of a full blown trial and appeal could no doubt be defeated more cheaply in the administrative context. In point of fact, opposition proceedings would harness the very market knowledge on which Lemley focuses. Market knowledge would simply be invoked *ex ante*, in the context of a low cost administrative proceeding, rather than *ex post*, in the context of a high cost judicial proceeding. Indeed, Lemley ultimately acknowledges that his analysis may be consistent with third-party oppositions and even notes data from Europe indicating that opposition proceedings are the single best predictor of a patent's value.²¹⁶

At the end of the day, then, Lemley is partly right. Turning the PTO into a full-fledged administrative agency, whose fact finding would be subject to deference in all contexts, would probably not be cost effective. By the same token, there is reason to believe that improvement through the mechanism of an opposition proceeding would be extremely cost effective. With an opposition proceeding, the more expensive artillery of a court proceeding would have to be deployed in a significantly smaller number of cases.

Like Mark Lemley, Ian Ayres and Paul Klemperer offer a probing economic critique of a system in which administrative agency proceedings provide early confidence about patent rights.²¹⁷ They assert that uncertainty about a patent's validity is actually valuable because it can reduce the deadweight loss associated with patent monopolies without significantly altering a patentee's profits (and hence its incentive to innovate).²¹⁸ The Ayres and Klemperer position runs as follows: Suppose courts made determinations about whether patents were valid by flipping weighted coins at the end of the patent's twenty-year life. If a patent was found valid—and the probability of such a finding would be high but less than one—the patentee would be entitled to compensatory damages from any parties that had infringed during the patent's life.²¹⁹ This type of uncertain situation would induce some level of infringement and hence reduce the price the patentee could charge. Reduction in the monopoly price would reduce both deadweight loss and patentee profits.

216. See *id.* at 1524–25 & n.111. Lemley also acknowledges that his analysis is consistent with proposals to change incentive structures in the PTO so that examiners are rewarded not only for granting patents but also for rejecting them. See *id.* at 1523.

217. See Ayres & Klemperer, *supra* note 196.

218. See *id.* at 987–89.

219. *Id.* at 994.

Significantly, however, because the last increment of monopoly pricing actually produces little additional revenue for the patentee, the reduction in deadweight loss would be much greater than the reduction in patentee profit.²²⁰ For example, in a simple linear model, if the probability of enforcement was ninety-five percent, the market price would drop ten percent. This ten percent drop in price would reduce deadweight loss by nineteen percent but would reduce the patentee's expected profit by only one percent.²²¹

The model put forward by Ayres and Klemperer is, however, quite unrealistic. As Ayres and Klemperer acknowledge, efficient reductions in deadweight loss occur only in the particular case where there is uncertainty about whether a valid patent will be upheld (what they term "Type I" uncertainty).²²² If the uncertainty runs in the opposite direction—in other words, where there is uncertainty about whether an invalid patent will be held valid—there is no deadweight loss reduction. To the contrary, the latter type of uncertainty—Type II uncertainty—*produces* unnecessary deadweight loss. The stylized model put forward by Ayres and Klemperer only admits of Type I uncertainty—that is, the possibility that valid patents will be held invalid.²²³ In the real world, however, if the PTO were to act largely as a registry of patents, the validity of which was not ultimately decided by the courts until after the twenty-year patent term had expired, both Type I *and* Type II uncertainty would be present. In addition, the Ayres and Klemperer model fails to take into account any role for patents other than their "classical" role in providing incentives to invent. To the extent that the security of a monopoly position in the area of a particular invention is necessary to spur further innovation in that area,²²⁴ uncertainty about a patent's validity will reduce investments for development. Ayres and Klemperer also assume risk neutrality.²²⁵ In situations where patents are held by large corporations that have other mechanisms for hedging risk, their assumption may be accurate. However, for small start-up firms, patents are often the only risk-hedging assets owned by the firm. Finally, and most relevantly for purposes of this article, if a reduction in the market power of patentees is in fact the best mechanism for achieving reductions in deadweight loss, such reduction

220. *Id.* at 1000.

221. *Id.* at 1000–01. Ayres and Klemperer also argue that even in cases where market price, and hence profits, were substantially diminished, patent term extensions could compensate the patentee while keeping deadweight loss low relative to the case of full monopoly. See *id.* at 1001–07.

222. See *id.* at 1019–20.

223. See *id.* at 995 n.20 (acknowledging this point).

224. Edmund Kitch has made an influential argument that patents on nascent invention serve as "prospects" to induce further development and commercialization in the area of the given invention. Kitch, *supra* note 10, at 267–76; see also F. Scott Kieff, Property Rights and Property Rules for Commercializing Inventions, 85 Minn. L. Rev. 697, 707 n.47 (2000) (discussing Kitch's theory of prospects).

225. Ayres & Klemperer, *supra* note 196, at 997.

can be achieved without introducing uncertainty. As Ayres and Klemperer point out, a regime of partial damages, wherein patentees are not fully compensated for infringement, would reduce patentees' market power.²²⁶

At bottom, then, certainty has significant value in the patent system. Additionally, a greater level of certainty can be achieved in a relatively cost effective fashion through the institution of opposition proceedings, and judicial deference to patents that issue after such proceedings.

B. *Deference to Trial Courts*

Thus far I have discussed the role of deference by courts (whether trial courts or the Federal Circuit) to PTO fact finding. In this section, I discuss what role, if any, deference should play in Federal Circuit review of trial court fact finding.

1. *General Considerations.* — The general rule that limits the scope of judicial review over lower court fact finding is often justified by the claim that lower courts are more experienced in fact finding than appellate courts. Moreover, as the Supreme Court has emphasized, this experience comes into play whether or not the fact finding turns on credibility determinations.²²⁷ In the case of trial versus appellate courts (as contrasted with the case of agency versus court), the institutional competence argument has a certain question-begging quality. There is nothing in the nature of a trial court that makes it more capable, as a structural matter, of fact finding than an appellate court. In addition, there is no constitutional or statutory bar against appellate courts finding facts, or even hearing witnesses.²²⁸ Indeed, in those cases where witness credibility was at issue, appellate courts could even hear the same witnesses who appeared

226. See *id.* at 1028–31. Ayres and Klemperer also do not account for the extent to which existing regimes of price discrimination reduce deadweight loss. See Rai, *Information Revolution*, *supra* note 154, at 202–09 (discussing price discrimination in the area of patented pharmaceuticals).

227. *Anderson v. Bessemer City*, 470 U.S. 564, 574–75 (1985). Some early cases had suggested that findings of fact that were not based on credibility determinations might be subject to a standard of review more exacting than clearly erroneous review. See, e.g., *Orvis v. Higgins*, 180 F.2d 537, 539–40 (2d Cir. 1950) (“If [a trial judge] decides a fact issue on written evidence alone, we are as able as he to determine credibility, and so we may disregard his finding.”). In *Bessemer City*, however, the Supreme Court made it clear that the “clearly erroneous” standard of review applied irrespective of whether the findings of fact were based on credibility determinations. 470 U.S. at 574. Moreover, Rule 52(a) was amended in 1983 to specify that the clearly erroneous standard applied to all findings of fact, including fact finding from documentary evidence. See Comm. on Rules of Practice and Procedure, Judicial Conference of the U.S., Preliminary Draft of Proposed Amendments to the Federal Rules of Civil Procedure 5 (1983). Prior to 1983, Rule 52(a) had stated: “Findings of fact shall not be set aside unless clearly erroneous, and due regard shall be given to the opportunity of the trial court to judge the credibility of the witnesses.” *Id.*

228. Benjamin, *supra* note 37, at 353.

before the trial court.²²⁹ In other words, if appellate courts were to institute procedures under which they routinely found facts based on oral or documentary evidence,²³⁰ they would presumably become expert in the finding of fact. Implicit in the assignment of primary responsibility for fact finding to the trial court, therefore, is the need to conserve scarce judicial resources. If the appellate court attempted to acquire the district court's knowledge of any given factual setting, that acquisition would come at great expense.²³¹ Relative to that expense, the benefits of having another fact finder are, in the vast majority of cases, likely to be small.²³²

Even in cases of law application, there is no need for de novo review unless the process of law application produces legal rules that are properly applicable in future cases.²³³ As Professor Randolph Jonakait has argued, when

the factors underlying the trial court's ruling are either so variable or so specific that they are unlikely to recur, the appellate court will be unable to announce a useful rule for future application. In such cases, courts do not have to be concerned with

229. Witness credibility is not likely, however, to be particularly important in many cases of scientific fact finding.

230. Of course, as matters currently stand, appellate courts do not have such procedures. Rather, the appellate presentation of a case is typically limited to briefs and a single oral argument.

231. See, e.g., *Pierce v. Underwood*, 487 U.S. 552, 560 (1988) (remarking on the expense of requiring the appellate court "to undertake the unaccustomed task of reviewing the entire record").

232. *Id.*; see also *Bessemer City*, 470 U.S. at 574–75 ("Duplication of the trial judge's efforts in the court of appeals would very likely contribute only negligibly to the accuracy of fact determination at a huge cost in diversion of judicial resources."). There are only a few factual contexts in which the expense of de novo review might be justified. One context might involve legislative facts (including general scientific principles) that are not case- and time-specific and hence are likely to be important in a variety of cases. Cf. Davis, *supra* note 21, § 12:3 ("Legislative facts do not usually concern the immediate parties but are the general facts which help the tribunal decide questions of law and policy and discretion."). Another context is where the relevant facts are likely to change repeatedly between the trial and appellate court decisions, such that the trial court's fact finding will inevitably be stale by the time the case gets to the appellate court. See Benjamin, *supra* note 37, at 287–88. Because the facts that are most relevant to patent cases are adjudicative facts tied to a particular point in time, neither of these scenarios is applicable to patent law. Fact finding in one patent case may be relevant for other cases if the same patent is litigated in more than one case. However, as discussed further *infra* note 245, this situation does not require de novo appellate review; rather, it can in significant part be addressed through principles of issue preclusion.

233. See, e.g., Randolph N. Jonakait, *The Standard of Appellate Review for Scientific Evidence: Beyond Joiner and Scheffer*, 32 U.C. Davis L. Rev. 289, 325–26 (1999) (noting that a major function of appellate courts is to formulate rules of general applicability and "to treat like situations alike"); Cynthia Lee, *A New "Sliding Scale of Deference" Approach to Abuse of Discretion: Appellate Review of District Court Departures Under the Federal Sentencing Guidelines*, 35 Am. Crim. L. Rev. 1, 34 (1997) ("De novo or nondeferential review is generally most effective when the issue being reviewed is capable of being generalized by a rule of broad application.").

treating like cases in a dissimilar manner . . . [and] deference to trial court decisions is appropriate²³⁴

Notably, in emphasizing the appellate court's primary responsibility for maintaining a coherent body of legal doctrine, commentators like Jonakait follow in the footsteps of Supreme Court opinions that emphasize this primary responsibility.²³⁵

In contrast with fact finding or law application that is unlikely to produce generalizable rules, de novo review of strictly legal determinations is necessary for uniformity. This is particularly the case because single-judge district courts may lack the perspective and opportunity for deliberation necessary for a broad perspective on developing uniform law.²³⁶

2. *Should the Federal Circuit Defer?* — The question to which we turn next is the extent to which the Federal Circuit should defer. It might be argued that, because the Federal Circuit is itself a specialized court, with expertise superior to that of the trial courts, it should not be subject to standard deference principles with respect to fact finding.²³⁷ This argument has some merit. However, the conclusion to be drawn is not that the Federal Circuit should act as a second trial court—and third fact finder, after the PTO and the trial court.²³⁸

In the first instance, because any expertise that the Federal Circuit has rests on its familiarity with patent law, and, relatedly, generalizable principles of patent law application, there is no reason to believe that it

234. Jonakait, *supra* note 233, at 326–27.

235. See Evan Tsen Lee, *Principled Decision Making and the Proper Role of Federal Appellate Courts: The Mixed Questions Conflict*, 64 S. Cal. L. Rev. 235, 263 (1991) (analyzing in this light the Supreme Court decisions in *Icicle Seafoods, Inc. v. Worthington*, 475 U.S. 709 (1986) and *Pullman-Standard v. Swint*, 456 U.S. 273 (1982)). As Lee points out, any “rule” that could have been devised to govern the Title VII challenge in the *Pullman-Standard* case would have been too specific to have any precedential value. *Id.* at 265. In that case, according to Lee, the rule devised would be the following:

Where union at plant in fully-segregated town whose highest-paying positions were monopolized by workers of one race and which had history of mysteriously appearing and disappearing departments, effect of which was to perpetuate said monopolization and segregation, traded twenty-four workers of first race for two workers of second race, resulting in entire union membership of first race, said union intentionally discriminated on basis of race within meaning of § 703(h).

Id.; see also *Salve Regina Coll. v. Russell*, 499 U.S. 225, 233 (1991) (suggesting deference by appellate courts “when it appears that . . . probing appellate scrutiny will not contribute to the clarity of legal doctrine”).

236. See, e.g., Lee, *supra* note 235, at 251 (noting that appellate judges have the “advantage of group deliberations and group voting” and “also have the luxury of distanced reflection”).

237. See, e.g., Dreyfuss, *Federal Circuit*, *supra* note 2, at 47–53 (making this argument).

238. See, e.g., *Markman v. Westview Instruments, Inc.*, 52 F.3d 967, 1025 (Fed. Cir. 1995) (Newman, J., dissenting) (“[I]t is an illusion to think that patent litigation difficulties can be resolved by turning factual issues into matters of law and assigning them to the Federal Circuit.”).

will be familiar with the factual particulars of any given technology. For this reason, as was discussed earlier, the Federal Circuit has performed poorly when finding facts de novo in biotechnology patent appeals from the PTO.²³⁹ Similarly, it is by no means clear that the Federal Circuit performs satisfactorily in its de novo review of trial court fact finding. For example, the Federal Circuit's tendency to believe that the level of ordinary skill in the art allows software programs to be enabled without disclosure of source code, flow charts, or other detail has led it to overturn enablement findings by lower courts that paid more careful attention to the actual expert testimony in the case.²⁴⁰

In addition, even if the Federal Circuit could do a perfect job of fact finding, de novo fact finding by an appellate court is hardly efficient. The resources of both the trial court and the appellate court must be expended on questions that will typically matter in only a single case. Moreover, there is delayed certainty about the patent right, as both parties know that all aspects of a district court's determination will be reviewed anew on appeal.²⁴¹ This inefficiency manifests itself with particular force in the context of de novo review of claim construction. Because both patent validity and infringement often turn on claim construction, such construction often determines the ultimate outcome of the case. Unless a district court's claim construction can properly form the basis for a summary judgment motion, however, the district court's claim construction will not be reviewed until after a full trial on validity or infringement has taken place.²⁴² If the claim construction is then reversed on appeal, the domino effect discussed earlier comes into play, and the issues of validity or infringement are likely to be called into question.²⁴³ In those cases where the appellate court does not do its own fact finding regarding validity or infringement,²⁴⁴ the case must be returned to the trial court for another trial. At bottom, even if de novo review of fact finding by the Federal Circuit meant a greater likelihood of achieving justice between the parties, this increased accuracy would come at great

239. See Burk & Lemley, *supra* note 132, at 1199–2000 (making similar point); *supra* notes 151–160 and accompanying text.

240. See *supra* text accompanying notes 127–132.

241. Perhaps not surprisingly, given the Federal Circuit's de novo review standard, the available evidence suggests that parties tend to see trial court decisionmaking as simply an opening step on the road to appeal. For example, as noted earlier, according to one survey, only twenty-nine percent of cases settle pursuant to a trial court claim construction. *Markman* Survey, *supra* note 48, at 15.

242. The Federal Circuit does not hear interlocutory appeals of claim construction. See *Cybor Corp. v. FAS Tech., Inc.*, 138 F.3d 1448, 1479 (Fed. Cir. 1998) (Newman, J., dissenting). For criticism of the Federal Circuit's refusal to hear interlocutory appeals, see Craig Allen Nard, *Process Considerations in the Age of Markman* and *Mantras*, 2001 U. Ill. L. Rev. 355, 356–57.

243. See *supra* text accompanying notes 108–109.

244. As noted earlier, the Federal Circuit does do such fact finding on occasion, particularly with respect to infringement. See *supra* note 109 and accompanying text.

expense and would typically do nothing to further the broader goal of legal uniformity.²⁴⁵

It could be argued that while routine fact finding by the Federal Circuit is inaccurate and inefficient, its efforts to create bright-line rules that obviate the need for trial court fact finding are more laudable. For example, as noted earlier, the Federal Circuit's biotechnology jurisprudence has reduced the need for inferior decisionmakers to look closely at the prior art or disclosure in individual cases.²⁴⁶ Similarly, in the area of computer technology, the court appears to be moving towards a bright-line rule regarding disclosure.²⁴⁷ The Federal Circuit's efforts to turn trial court claim construction into a strictly legal exercise guided by the four corners of the patent document might also be seen as a move toward rule-oriented decisionmaking that increases predictability and reduces trial court discretion. Again, however, an embrace of bright-line rules over case-specific factual inquiry assumes that the initial allocation of rights is largely irrelevant, so long as it is predictable.²⁴⁸ Further, at least in the context of claim construction, any attempt to seek a high level of predictability may be illusory. Despite the Federal Circuit's efforts over the last few years, its reversal rates with respect to lower court claim con-

245. Admittedly, there is one situation where deferential review of lower court fact finding could create legal inconsistency. This situation might arise where the same patent was asserted against different parties in different trial courts. These trial courts might, based on divergent factual findings, reach divergent conclusions on the patent's validity. If the appellate court were to defer to the factual findings of the trial court in each case, it would not be able to reconcile the divergent legal conclusions. For two reasons, however, this situation should not arise frequently. As an initial matter, at least for purposes of patents that have been through opposition proceedings, the factual findings of trial courts should be similar. As noted earlier, for patents that have been through opposition proceedings, there are sound institutional reasons for trial court deference with respect to factual issues argued before the PTO. In addition, even where there was diminished trial court deference due to the absence of an opposition proceeding, the possibility of legal inconsistency could be addressed through principles of issue preclusion. Under the doctrine of non-mutual defensive issue preclusion, when one court has decided a question of law or fact against a patentee, future defendants are allowed to assert the prior court's finding against the patentee. Indeed, the seminal Supreme Court case on issue preclusion, *Blonder-Tongue Laboratories, Inc. v. University of Illinois Foundation*, 402 U.S. 313 (1971), was a case involving precisely this type of preclusion. The *Blonder-Tongue* Court grounded its preclusion decision on policy concerns about preservation of judicial resources and diversion of litigant resources from alternate productive uses. *Id.* at 328-29. Of course, even with defensive issue preclusion, the possibility of legal inconsistency with respect to a finding that had initially been decided in the patentee's *favor* would still exist. But to the extent that there were a single specialized trial court (one of the options I advocate below, see *infra* Part II.B.2.d.) even this residual category of cases would be eliminated.

246. See *supra* text accompanying note 173.

247. See Burk & Lemley, *supra* note 132, at 1162-64 (discussing elimination of disclosure requirements in software and noting that "[i]t is simply unrealistic to think that one of ordinary skill in the programming field can necessarily reconstruct a computer program given no more than the purpose the program is to perform").

248. See *supra* note 174 and accompanying text.

struction are still quite high. Either trial courts are unable to follow the Federal Circuit's bright-line directions, or these bright-line directions are simply ill-suited to claim construction.

Rather than looking to appellate fact finding or bright-line rules that obviate the need for fact finding, we should bolster the expertise of trial courts. In what follows, I make the argument for why such bolstering is necessary. I first argue that deference to fact finding by the PTO (on the part of either the trial court or the appellate court) will not be appropriate in all cases. Hence a second layer of fact-finding expertise at the trial court is necessary. I then review evidence, both anecdotal and statistical, that underscores the limitations of jury fact finding. Based on this evidence, I make suggestions for cabining the role of the jury. I also address the competence of generalist trial court judges in finding facts, arguing that while their competence may be superior to that of the jury, there is reason to believe that specialized trial courts would be useful.

a. *Why Deference to the PTO is Insufficient.* — One might question the need for fact-finding expertise, at least with respect to questions of patent validity (as contrasted with patent infringement), at any level within the court system. On this view, if the PTO were given sufficient fact-finding expertise, the trial court and the appellate court could simply defer to such expertise. As discussed earlier, however, there are reasons to question deploying the PTO's resources in its review of all patent applications. Rather, the most cost effective way to deploy these resources would be selectively, through some type of opposition proceeding. As a consequence, in at least some cases where patents are litigated, the patent will not have been examined rigorously. This will be particularly true in situations where the economic significance of a particular patent, and hence the need to oppose it, is not fully understood until several years after it is issued. In those cases, significant deference to the PTO's fact finding will not be appropriate.²⁴⁹ Moreover, even where the patent has gone through an opposition proceeding, there may be cases where the proceeding does not bring forward some of the factual arguments that are subsequently brought before the trial court.

As matters currently stand, however, there are reasons to have serious concern about the competence of the trial court. Limits on competence are particularly well documented in the case of the trial court jury.

b. *The Limitations of the Trial Court Jury.* — At the outset, it is important to emphasize that the problem posed by jury trials has become more acute in the last decade, as the role of juries has increased dramatically. Early on, juries were not heavily involved in patent cases. Even as late as 1970, only 2.6% of all cases were tried to juries.²⁵⁰ By the 1990s, however,

249. For this reason, the approach taken in the European system, where the European Patent Office has plenary power over validity judgments, is probably not appropriate. See *supra* text accompanying note 190.

250. Moore, Black Box, *supra* note 111, at 366 n.6.

a trend toward jury trials was evident. On average, in the years 1997 to 1999, fifty-nine percent of cases were tried to juries.²⁵¹

As the prevalence of juries in patent cases has increased, patent law practitioners and scholars have increasingly questioned jury competence. Those who have complained stress jurors' limited comprehension of the difficult science and technology at issue in many patent cases;²⁵² the tendency of juries to favor the patentee;²⁵³ and their tendency to be swayed by irrelevant factors, such as the identity of the plaintiff and defendant.²⁵⁴ Although these anecdotal complaints are hardly dispositive, they are supported to some extent by more general empirical studies of how well jurors understand complex cases.²⁵⁵

To be sure, these general empirical studies do not focus on patent cases. Moreover, there are other studies suggesting that jurors do a reasonably good job in understanding the evidence in complex cases.²⁵⁶ The best statistical evidence of the limitations of trial court jurors in the specific case of patent law is provided by Kimberly Moore's analysis of the 1,209 patent cases that were resolved by a trial court fact finder (whether judge or jury) between 1983 and 1999.²⁵⁷ Moore's examination of these cases indicates that, where a patentee brings the lawsuit, juries rule for patentees in sixty-eight percent of cases.²⁵⁸ By contrast, judges rule for patentees in only fifty-one percent of cases.²⁵⁹ The discrepancy between the behavior of judges and juries is particularly striking to the extent one can assume that defendants have incorporated information about jury

251. *Id.* at 366 n.7.

252. *Id.* at 365.

253. Allan N. Littman, *The Jury's Role in Determining Key Issues in Patent Cases: Markman, Hilton Davis and Beyond*, 37 *IDEA* 207, 209 (1997).

254. See Fourth Biennial Patent System Major Problems Conference, *Abolition of Jury Trials in Patent Cases*, 34 *IDEA* 77, 87-88 (1994) (statement of patent lawyer Dick Witte). Witte observes that post-trial interviews reveal that

juries are much more susceptible than our judges to being influenced by irrelevant issues. These are issues not directed toward the merits of infringement or validity, such as the large corporation and the small corporation syndrome, just as an example. Juries are also more susceptible to being influenced by demagoguery by lawyers, particularly that dwell on such issues as the horrors of injunctive relief or the bad aspects of patent as a monopoly.

Id.

255. See Robert MacCoun, *Inside the Black Box: What Empirical Research Tells Us About Decisionmaking by Civil Juries*, in *Verdict: Assessing the Civil Jury System* 137, 151 (Robert E. Litan ed., 1993) (noting that a "growing body of studies indicates that some jurors may fail to comprehend as much as 50 percent of the judge's instructions").

256. See, e.g., Marc Galanter, *The Regulatory Function of the Civil Jury*, in *Verdict: Assessing the Civil Jury System*, supra note 255, at 61, 70 ("The literature, on the whole, converges on the judgment that juries are fine decisionmakers. They are conscientious, collectively they understand and recall the evidence as well as judges, and they decide on the basis of the evidence presented.").

257. Moore, *Black Box*, supra note 111, at 386.

258. *Id.*

259. *Id.*

bias into their decision to take the case to trial in the first instance. Assuming this information has been incorporated, even if juries are biased towards patentees, win rates before juries should approximate win rates before judges.²⁶⁰ Moore's data therefore indicate not only that juries are more favorably disposed towards patentees than judges but also that infringement defendants may routinely underestimate the extent of jury bias. Such faulty estimation may be a consequence of the fact that juries have only recently become a common feature of patent cases.²⁶¹ It is also likely, however, that jury decisions will remain intrinsically less predictable than judge decisions. While judges can develop a reputation in the patent trial arena, particularly because they write opinions enunciating findings of fact and conclusions of law, juries are one-shot players that often deliver "black box" decisions.

Other data from Moore's study suggest that juries are less likely than judges to understand the nuances of patent cases. For example, on the independent issues of patent validity and infringement, juries find for the same party eighty-six percent of the time; by contrast, judges find for the same party only seventy-four percent of the time.²⁶² These data suggest that judges are more likely than juries to understand that validity and infringement are not inextricably linked. Even more ominously, jury decisions heavily favor whichever party filed the suit. As noted earlier, in the ordinary infringement case, where the patentee files the lawsuit, juries rule for the patentee on sixty-eight percent of litigated claims; by contrast, in declaratory judgment actions initiated by a potential infringer, the patentee wins only thirty-eight percent of the time.²⁶³ Unlike juries, judges tend to rule for the patentee in approximately the same percentage of claims irrespective of who files the lawsuit.²⁶⁴

Finally, as the Federal Circuit is fond of noting, it is very difficult for courts to engage in meaningful review of "black box" jury verdicts. This is

260. Moreover, according to the fifty percent rule propounded by Priest and Klein, both of these rates should stand at about fifty percent, in that the only cases that should go to trial are cases where it is genuinely difficult to ascertain which way the fact finder will rule. See Priest & Klein, *supra* note 104 at 4-5; *supra* text accompanying notes 104-106. It bears mention, however, that the assumption regarding symmetric stakes that underlies the Priest and Klein model does not necessarily apply to patent trials. As a general matter, the patentee is likely to have more to lose than the infringer; if the patent is invalidated, this invalidation applies not only against the infringer but also against the world at large. Thus the patentee may be more willing to settle close cases than to let them go to trial. See Daniel Kessler et al., *Explaining Deviations from the Fifty-Percent Rule: A Multimodal Approach to the Selection of Cases for Litigation*, 25 J. Legal Stud. 233, 237-42 (1996) (noting likelihood of divergence from fifty percent rule in cases where litigants have differential stakes).

261. Moore, *Black Box*, *supra* note 111, at 366 (noting that while under three percent of patent cases tried in district court went to juries between 1968 and 1970, fifty-nine percent of such cases went to juries between 1997 and 1999).

262. *Id.* at 403 & fig.10.

263. *Id.* at 406 fig.12.

264. *Id.*

a particular problem when it comes to the central question of patent validity. Dissenting in the 1995 case *In re Lockwood*, the late Judge Nies observed:

As jury cases are now tried, in accordance with our precedent, the evidence respecting validity of a patent is thrown into the black box of the jury room, and the verdict is returned either valid or invalid. If both parties agree to that procedure, so be it. But where a party objects, I believe that a litigant has a right to a trial court's decision with findings of fact and conclusions of law on the issue of validity.²⁶⁵

c. *Cabining the Role of the Patent Jury*. — Given the shortcomings of trial court juries, one mechanism for improving trial court expertise might involve limiting the role of the jury, at least with respect to technical patentability questions that do not require value judgments or judgments regarding credibility. Proposals to limit the role of the jury raise, of course, serious Seventh Amendment concerns. Although a full discussion of the arguments for and against having juries in technically complex civil cases is beyond the scope of this Article, a few observations are in order. First, as a doctrinal matter, the Supreme Court has indicated that functional analysis is often applicable in patent cases and can be used to cabin the role of the jury in deciding technical questions. Second, the Supreme Court's decision to emphasize functional considerations is hardly ungrounded: To the extent that the Framers thought the jury to be necessary for purposes of infusing the law with popular values,²⁶⁶ that role is not undermined by taking away from the jury technical questions of patentability.

The Court's test for determining application of the Seventh Amendment looks initially to whether the cause of action in question, or actions similar to the cause of action, were tried to juries prior to 1791.²⁶⁷ In the patent context, the outcome of this initial examination is clear. As the Supreme Court noted in its *Markman* decision, because infringement cases were clearly tried to juries in the eighteenth century, the jury trial

265. 50 F.3d 966, 990 (Fed. Cir. 1995) (Nies, J., dissenting).

266. See, e.g., Leon Green, *Jury Trial and Mr. Justice Black*, 65 Yale L.J. 482, 483 (1956) (noting that the jury "offers an assurance of judgment by neighbors who understand the community climate of values"); Charles W. Wolfram, *The Constitutional History of the Seventh Amendment*, 57 Minn. L. Rev. 639, 671-72 (1973) (arguing that the Anti-Federalists who fought for the Seventh Amendment were concerned with the danger of placing control in the hands of elite judges).

267. See, e.g., *Pernell v. Southall Realty*, 416 U.S. 363, 370-76 (1974) (addressing issue of jury trial in statutory eviction action); *Curtis v. Loether*, 415 U.S. 189, 192-98 (1974) (addressing issue of jury trial under Civil Rights Act of 1968). Though the historical focus of the Seventh Amendment test is longstanding, it has been criticized. See, e.g., Matthew P. Harrington, *The Economic Origins of the Seventh Amendment*, 87 Iowa L. Rev. 145, 148 (2001) (arguing that tying the scope of the jury trial right to pre-1791 case law misunderstands the history of the Seventh Amendment).

right attaches to patent cases.²⁶⁸ The extent to which the jury trial decision attaches to any *particular decision* in a patent case is much less clear, however. Indeed, the Court's reasoning in *Markman* suggests that although the Seventh Amendment would prevent the elimination of the jury in patent cases, it would not pose a significant impediment to a reduction in the jury's role. In *Markman*, the Supreme Court argued that although the jury trial right attaches to patent cases, it attaches to specific questions of patentability only to the extent that juries decided analogues to those questions in eighteenth-century practice. In the case of claim construction, there was no eighteenth-century analogue.²⁶⁹

Once the *Markman* Court had determined that the Seventh Amendment's historical test did not provide an answer to the allocation of power question, it used explicitly functional considerations to decide the question. The Court noted that claim construction is

a special occupation, requiring . . . special training and practice. The judge, from his training and discipline, is more likely to give a proper interpretation to such instruments than a jury; and he is, therefore, more likely to be right, in performing such a duty, than a jury can be expect to be.²⁷⁰

The Court also emphasized that a jury's capacity for evaluating demeanor, sensing the "mainsprings of human conduct," or reflecting community standards is much less important in evaluating scientific facts in patent cases than in evaluating other types of facts.²⁷¹

To the extent that the *Markman* reasoning is followed in subsequent cases, not only will claim construction be an issue removed from the jury, but so too will any technical patent question not specifically analogous to an issue decided by a jury in the eighteenth century. To be sure, *Markman's* approach to the Seventh Amendment diverges from the traditional approach the Court uses in deciding Seventh Amendment questions. Traditionally, once the Court has made the initial determination that a jury trial right attaches, it follows the fact/law distinction in allocat-

268. *Markman v. Westview Instruments, Inc.*, 517 U.S. 370, 377 (1996) (noting the "descent of today's patent infringement action from the infringement actions tried at law in the 18th century" and the lack of any dispute that "infringement cases today must be tried to a jury, as their predecessors were more than two centuries ago").

269. *Id.* at 379–84.

270. *Id.* at 388–89 (quoting *Miller v. Fenton*, 474 U.S. 104, 114 (1985)).

271. *Id.* at 389–90. As David Faigman explains in regard to the usefulness of demeanor,

[g]ood scientific research simply does not depend on the credibility of individual witnesses. If the question is whether the declarant made a statement under a belief of impending death, the nurse's credibility might be critical In contrast, whether a series of six epidemiological studies support the conclusion that the relative risk associated with silicone implants exceeds 2.0 for connective tissue disorders does not entail the same sort of credibility assessment.

David L. Faigman, Appellate Review of Scientific Evidence Under *Daubert* and *Joiner*, 48 Hastings L.J. 969, 978–79 (1997).

ing power between jury and judge.²⁷² Nonetheless, there is every reason to believe that *Markman*'s decision to decouple the fact/law distinction from the determination of how to allocate power between judge and jury will be followed in subsequent patent law cases.

The Court's doctrinal decision to deviate from its traditional emphasis on the fact/law distinction and to rely instead on functional considerations can also be defended normatively. It is by no means clear that the fact/law distinction on which the Court typically relies tracks well the concerns that animated proponents of the Seventh Amendment. To the contrary, where factual decisions are primarily technical in nature, they do not trigger the concerns about oppressive rule by elites that the Seventh Amendment was intended to address.²⁷³ The Supreme Court's emphasis on functional considerations in deciding the allocation of power between judge and jury also has a substantial pedigree within the common law. In many areas of the common law, the key determinant of what the jury must decide is not where a question falls on the fact/law spectrum. Rather, it is the extent to which the answer to the question is thought to turn on community values. Thus, as Mark Gergen has noted, in negligence cases, juries often decide questions of reasonableness even when facts are not in serious dispute.²⁷⁴ By contrast, in contract law, where economic considerations are paramount, juries have tended to have a more limited role, even in deciding factual questions.²⁷⁵

One might argue that "blue ribbon" juries composed of individuals who were technically trained might address concerns about the technical competence of juries without impinging on the jury's traditional role as fact finder. This suggestion is unlikely to be feasible, however. Because technical expertise in a particular area is not likely to be useful if the patent covers an area unrelated to that expertise, simply assembling technically competent jurors is not sufficient. Moreover, assembling a group of jurors trained in the area of the patent at issue would, in all likelihood, be prohibitively expensive.

For these reasons, the concerns about juries deciding technical questions that animated the Supreme Court decision in *Markman* should be

272. See, e.g., *Colgrove v. Battin*, 413 U.S. 149, 157 (1973) (noting that a district court rule permitting six rather than twelve persons to constitute a jury did not violate the Seventh Amendment because it did not impact the jury's duty of "assur[ing] a fair and equitable resolution of factual issues"). As noted above, in *Markman*, the Supreme Court adhered only to the first part of its traditional test. See *supra* text accompanying note 268.

273. The long-mooted and highly controversial question of whether there might exist some type of narrow "complexity" exception to the Seventh Amendment, see Graham C. Lilly, *The Decline of the American Jury*, 72 U. Colo. L. Rev. 53, 80-81 (2001) (discussing Supreme Court's suggestion of this exception in *Ross v. Bernhard*, 396 U.S. 531 (1970), and Third Circuit's application of this exception), might also be defended on the grounds that such complex questions typically do not turn on popular value judgments. Traditionally, the complexity exception has been defended on due process grounds.

274. Gergen, *supra* note 138, at 410.

275. See *id.* at 411-12.

heeded. Consistent with the Seventh Amendment, it should be possible to take questions regarding patent validity and infringement that are purely technical in nature away from the jury. In many cases, this would mean that a jury would simply not be involved in patent lawsuits.

d. *Trial Court Judges: The Case for Specialization.* — The limitations of the trial court jury raise the question of how qualified generalist trial court judges are to address patent cases. As was noted previously, the evidence regarding trial court judges does suggest some points in their favor.²⁷⁶ Judges are not swayed by the variable of who brought the lawsuit. Judges are also better able than juries to distinguish between the independent questions of patent validity and patent infringement. Moreover, unlike juries, judges write opinions that have findings of fact and conclusions of law. They do not render “black box” verdicts.

By the same token, generalist trial judges also have significant limitations. The infrequent nature of patent litigation makes it unlikely that the typical district judge will see more than a few patent cases over the course of her tenure.²⁷⁷ This is particularly the case if the judge is outside one of the relatively small number of jurisdictions that have a disproportionate percentage of the total patent litigation case load.²⁷⁸ As a consequence, many trial court judges may be uncomfortable dealing with the intricacies of patent litigation.

One mechanism for improving the expertise of patent trial courts might involve the establishment of specialized trial courts. To some extent, we are edging towards having such courts *de facto*, even independent of any specific legislative reform to implement such courts. For example, the ten district courts that have the highest number of patent cases also hear about forty-four percent of all patent cases.²⁷⁹

Nonetheless, the majority of cases (fifty-six percent) still go to inexperienced courts. Even if this percentage were to decrease somewhat, it would probably be far from zero: When parties can choose between trial tribunals with expertise in the intricacies of patent litigation, and those without such expertise, at least some parties with weak cases may gamble on non-expert tribunals.²⁸⁰ The current system is also problematic in that the Federal Circuit does not appear to give fact finding by more experienced district courts substantially greater weight than fact finding by less experienced ones. Data collected in a recent study by Christian Chu show no statistically significant difference in reversal rates between “more

276. See *supra* text accompanying notes 250–265.

277. One commentator estimates that the average judge has “only one patent trial every 6 to 8 years.” John B. Pegram, *Should There Be a U.S. Trial Court with a Specialization in Patent Litigation?*, 82 J. Pat. & Trademark Off. Soc’y 765, 788 (2000).

278. See Moore, *Forum Shopping*, *supra* note 194, at 903–04 (noting that ten district courts resolve forty-four percent of all patent cases).

279. *Id.*

280. Cf. *id.* at 932–37 (discussing reasons for forum shopping at trial level).

active" patent tribunals and "less active" ones.²⁸¹ In contrast with our current system of semi-specialization, a specialized trial court would have the imprimatur of authority with respect to fact finding. This added authority would presumably add to the deference given by the Federal Circuit.

One important question regarding the design of a specialized district court involves the issue of whether judges on the court should be scientific specialists. Scott Brewer has forcefully argued that only judges with training in the science or technology involved in a particular case have the "epistemic competence" to adjudicate disputes in scientific or technological areas.²⁸² Brewer's proposal for judges who can wear the hats of both legal decisionmaker and scientist is important and provocative. By showing us what strict adherence to an ideal of judicial competence in technically complicated cases might look like, Brewer demonstrates how far we have to travel. One obvious problem with his proposal is feasibility.²⁸³ Because expertise in one area of science and technology does not transfer over to other areas,²⁸⁴ application of Brewer's proposal to the patent system would require a trial court with as many specialties and subspecialties as the PTO. It is unlikely that we will be able to assemble a group of judges who not only are sophisticated in the law but also have expertise in all of the various areas of science and technology covered by the patent system. Even if we could, setting up a judicial process that is merely a higher cost version of the administrative process is unlikely to produce benefits that would justify its cost.

By contrast, it would not be difficult to assemble a group of judges who had some exposure to basic scientific research methodology. These judges could then be given a budget sufficient to encourage liberal use of court-appointed expert witnesses, special masters, and technical advisors who had skill in the particular science or technology relevant to the patent case.²⁸⁵ To be sure, liberal use of such third-party expertise would

281. Chu, *supra* note 102, at 1121–23. As noted earlier, Chu analyzed patent decisions rendered by the Federal Circuit between January 1, 1998 and April 30, 2000. *Id.* at 1092. The category of more active tribunals "included district courts from which the Federal Circuit reviewed more than ten cases during the studied period." *Id.* at 1122. The more active group also included tribunals with specialized jurisdiction that includes patents, such as the Board of Patent Appeals and Interferences, the Court of Federal Claims, and the International Trade Court. *Id.* Chu's analysis did find that, within his studied population, the Federal Circuit tended to reverse more cases from less active tribunals than from more active ones (forty-one versus thirty-four percent, respectively). *Id.* at 1123. This tendency was not sufficiently marked, however, to have predictive implications for future cases. *Id.*

282. See Scott Brewer, *Scientific Expert Testimony and Intellectual Due Process*, 107 *Yale L.J.* 1535, 1565 (1998).

283. I address Brewer's argument more fully in Arti K. Rai, *Specialized Trial Courts: Concentrating Expertise on Fact*, 17 *Berkeley Tech. L.J.* 877, 879–80 (2002).

284. See Kesan, *supra* note 149, at 766.

285. The Supreme Court, in particular Justice Breyer, has encouraged trial courts to take steps along these lines in technically complex cases. See *Gen. Elec. Co. v. Joiner*, 522 U.S. 136, 149 (1997) (Breyer, J., concurring); see also Frederick Schauer, *The Dilemma of*

tend to move aspects of patent litigation away from an adversarial model. It is by no means clear, however, that adversarial procedures that rely on a “battle of the experts” represent the best mechanism for educating lay persons about the relevant science.²⁸⁶

Admittedly, because specialized courts are prone to certain institutional pathologies, such as bias²⁸⁷ and tunnel vision,²⁸⁸ we should be cautious about creating such courts. Indeed, in Part III, I discuss the extent to which the Federal Circuit has fallen prey to these pathologies. One mechanism for addressing the perils of specialization would be to designate a number of different district courts as venues for patent cases (for example, one district court per circuit): Such dispersion of responsibility would presumably allow each district court to maintain a docket of non-patent cases. Even with such dispersion of responsibility, however, bias and tunnel vision are possible, particularly if the other cases before the court do not force the court to consider arguments regarding competition or economic policy more generally. As a consequence, as discussed below, the Federal Circuit may be affected by bias and tunnel vision even though patent cases make up only about twenty percent of its docket. This is because the other cases heard by the Federal Circuit do not force it to engage broad questions of economic policy. Ultimately, the better answer involves recognizing that even though bias and tunnel vision are possible, these conditions are less likely to be problematic at the trial level than at the appellate level.

Consider first the problem of bias. On initial examination, bias might appear to be a significant problem for any court that hears large numbers of patent cases but does not regularly hear, for example, anti-trust cases. Because the major firms in most industries are now patent

Ignorance: *PGA Tour, Inc. v. Casey Martin*, 2001 Sup. Ct. Rev. 267, 285 (2001) (discussing Justice Breyer’s interest in increasing the level of fact-finding expertise in the court system).

286. See Brewer, *supra* note 282, at 1619–25 (noting that when lay individuals are evaluating technically challenging questions, the usual mechanisms for deciding whether to give weight to particular testimony—for example, the demeanor of the witness—are not likely to be very useful).

287. See, e.g., Harold H. Bruff, *Specialized Courts in Administrative Law*, 43 Admin. L. Rev. 329, 331–32 (1991) (suggesting bias from capture of appointment process and second-guessing of agencies by specialized courts); Dreyfuss, *Federal Circuit*, *supra* note 2, at 3 (observing that specialized judges may be “susceptible to ‘capture’ by the bar that regularly practices before them”); Richard L. Revesz, *Specialized Courts and the Administrative Lawmaking System*, 138 U. Pa. L. Rev. 1111, 1146 (1990) (“[S]pecialized courts tend to [control administrative action] less effectively than generalist courts because they are more likely to exhibit systemic biases, often in the agency’s favor.”).

288. As many scholars have noted, sound decisionmaking on legal issues results from exposure to a wide range of legal problems. See, e.g., Bruff, *supra* note 287, at 331 (“A primary cost of specialization is loss of the generalist perspective.”); Revesz, *supra* note 287, at 1120 (“[T]he disadvantage of inexperience is often more than made up for by the advantage of a fresh outlook and broad viewpoint.” (quoting Justice Antonin Scalia, Remarks Before the Fellows of the American Bar Foundation and the National Conference of Bar Presidents 9–10 (Feb. 15, 1987))).

holders,²⁸⁹ the patent lawyers who represent these firms will be wary of making arguments to the court that significantly undermine overall patent protection: Even the lawyers who represent the defendant in a given case will tend to restrict themselves to factual, case-specific arguments (for example, arguments about prior art) that undermine the plaintiff's particular patents. In the context of the trial court, however, we need not be overly concerned about whether the court hears the full range of legal or policy arguments in favor of diminished patent protection. As discussed further below, the appellate court should have primary responsibility for policy development.²⁹⁰ Consequently, so long as the trial court is able to apply those arguments that it does hear to particular factual scenarios, it will be able to do a reasonable job of adjudication.²⁹¹ Relatedly, the loss of wider legal perspective that may result from specialization is much less likely to be a problem for a trial court that is focused on discrete questions of fact and law/policy application than for an appellate court that has broad responsibility for law/policy elaboration. Finally, as discussed further below,²⁹² because I argue for significantly greater generalist review at the appellate level, any net increase in specialization would be tempered through a more generalist appellate process.

The case for having specialized patent trial courts is bolstered by the success of specialized patent trial courts in other countries. For example, unlike the United States, the United Kingdom has chosen to institute specialization at the trial court level.²⁹³ Review of these trial court decisions is then conducted by generalist courts. Specialized trial courts with national jurisdiction have also enjoyed success in the domestic arena. For example, despite the availability of generalist district courts, litigants choose to bring the vast majority of tax cases to the specialist United States Tax Court.²⁹⁴ Although the Tax Court is headquartered in Washington, D.C., its judges ride circuit so as to reduce the inconvenience to taxpayers. The decisions of the Tax Court are then reviewed by generalist regional circuit courts.²⁹⁵ Similarly, if a specialized patent trial court were headquartered in one location, the judges of the court could ride circuit. Alternatively, if we had a number of specialized trial courts, a

289. See *supra* notes 176–177 and accompanying text.

290. See *infra* Part III.

291. See *infra* Part I.C.

292. *Infra* Part III.D.

293. See John F. Duffy, *Improving the Legal Process of Claim Interpretation: Administrative Alternatives*, 2 Wash. U. J.L. & Pol'y 109, 152–53 (2000) (discussing United Kingdom Patents Court, which hears all appeals from patent agency decisions and tries infringement cases).

294. In recent years, the percentage of tax cases tried in the tax court has reached ninety percent or more. See Sean A. Bryan et al., *Special Project, An Empirical Study of Intercircuit Conflicts of Federal Income Tax Issues*, 9 Va. Tax Rev. 125, 125 n.1 (1989); David Shores, *Rethinking Deferential Review of Tax Court Decisions*, 53 Tax Law. 35, 36 (1999) (noting that tax court is the forum of choice for most taxpayers).

295. See Shores, *supra* note 294, at 36.

convenient venue could be determined by the principles of the general corporate venue statute.²⁹⁶

Once fact-finding expertise were in place at both the PTO and trial court levels, the Federal Circuit's current refusal to defer to the fact finding of inferior decisionmakers, as well as its attempts to impose bright-line rules on factually complex patent inquiries, would clearly be unjustifiable. So even if the Federal Circuit did not move away from these strategies on its own, the Supreme Court would have every right to insist on the importance of fact finding by inferior decisionmakers. Indeed, in the context of appeals from PTO patent denials, the Court has already so insisted.²⁹⁷

C. The Relationship Between Fact Finding and Policy Application

Thus far this Article has argued against formal, bright-line rules on the grounds that the fact-intensive nature of the standards for determining patent scope and validity makes it very difficult to establish such rules. Indeed, in order for the Federal Circuit to set up bright-line rules in such areas as nonobviousness and adequate disclosure, it must ignore entirely the case- and time-specific realities of scientific innovation. To the extent that the fact-finding expertise of the inferior decisionmakers can be fortified, rules are an inappropriate mechanism for determining patent scope and validity.

The standard-filled language of the patent statute does not merely invite close analysis of the facts. It also invites judicial policymaking.²⁹⁸ In contrast, as discussed in Part III, the Federal Circuit's jurisprudence generally ignores policy considerations. Indeed, it is the court's disregard for policy that makes its jurisprudence particularly subject to characterization as formalism, specifically rule-formalism.²⁹⁹ Although rule-formalism encompasses decisionmaking that abstracts from factual complexity,³⁰⁰ most commentators stress its use as a mechanism for making

296. 28 U.S.C. § 1391(c) (2000).

297. See *supra* notes 80–85 and accompanying text. Given the Federal Circuit's resistance to *Dickinson v. Zurko*, 527 U.S. 150 (1999), however, another reminder may be necessary. See *supra* notes 86–90 and accompanying text.

298. For this reason, I will use the terms "patent law" and "patent policy" somewhat interchangeably throughout Part III. Although law and policy may represent different undertakings in some contexts, it is very difficult to differentiate law from policy in the patent arena. Cf. Sunstein, *supra* note 139, at 2086 (noting that resolving ambiguities in a statute is often a question of policy as much as it is one of law).

299. The precise scope of the more general term formalism is the subject of some dispute. See, e.g., Richard H. Pildes, *Forms of Formalism*, 66 U. Chi. L. Rev. 607, 607 (1999) (discussing formalism as "anticonsequentialist morality," "apurposive rule following," and "a regulatory tool for producing optimally efficient mixes of law and norms in contract enforcement regimes").

300. See Larry Alexander & Ken Kress, *Against Legal Principles*, 82 Iowa L. Rev. 739, 740 (1997) ("Rules are legal norms that are . . . triggered by a few easily identified factual matters.").

decisions independent of policy considerations.³⁰¹ Notably, however, just like avoidance of fact, the Federal Circuit's avoidance of policy could be justified by pointing to the infirmities of inferior decisionmakers. Many, if not most, policy-based directives are likely to require fact finding. For example, almost all of the policy-based directives discussed in the next Part (for example, the Supreme Court's "physicality" requirement for algorithm patents; a policy-based approach to business method patents; a rebuttable presumption in favor of a patentee's refusal to license its patent; the Supreme Court's recent decision on the scope of the doctrine of equivalents; giving narrower scope to patents on upstream research platforms)³⁰² require fact finding. The specific types of fact that will need to be found in each context are discussed further below. For present purposes, it suffices to say that if inferior decisionmakers lack expertise in fact finding, they will be unable to apply these directives properly. Conversely, to the extent that the fact-finding expertise of inferior decisionmakers is bolstered, there remains little justification for decisionmaking that abjures policy considerations.

In the next Part, I discuss the Federal Circuit's abstention from policy questions as well as reasons why appellate courts that address patent cases should in fact implement policy-based decisionmaking.

III. FILLING THE POLICY GAP

In this Part, I argue that although rule-formalism that is opaque to policy considerations may be a viable approach to judicial decisionmaking in certain areas, such formalism is a poor fit for a patent statute that is designed to evolve through policy-oriented judicial elaboration. The central difficulty with courts adopting policymaking roles, even when they are effectively delegated power to do so by Congress, is their limited resources and technical competence. Nonetheless, in the area of intellectual property, the appellate courts are probably superior as a policymaking body to the other institutional alternatives.

To be sure, a *specialized* appellate court—with its demonstrated commitment to formalism and clear potential for bias—may not provide the optimal framework for judicial policymaking. Therefore I discuss whether the best solution would involve abolishing the Federal Circuit, and having a system of specialized trial courts reviewed by generalist appellate courts. Although abolishing the Federal Circuit is most likely a political nonstarter, the Federal Circuit could, and should, be subject to sustained generalist review that forces it to evaluate, in a serious and neu-

301. See, e.g. Larry Alexander, "With me, It's All er Nuthin'": Formalism in Law and Morality, 66 U. Chi. L. Rev. 530, 531 (1999) (arguing that formalism means "adherence to a norm's prescription without regard to the background reasons the norm is meant to serve"); Daniel Farber, The Inevitability of Practical Reason: Statutes, Formalism, and the Rule of Law, 45 Vand. L. Rev. 533, 539–40 (1992) (discussing the formalist belief that judges should "follow the rules laid down").

302. See *infra* Part III.A; *infra* note 426 and accompanying text.

tral fashion, questions of innovation policy. Thus we should welcome the Supreme Court's recent interest in reviewing patent cases that raise not only allocation of power questions but also questions of substantive patent law and policy. The combination of serious generalist review and more trustworthy inferior decisionmakers should create a system where policymaking by the Federal Circuit is both possible and justifiable.

A. *The Federal Circuit: A Study in Formalism*

Rule-formalism that is opaque to policy considerations is often associated with textualist theories of statutory interpretation. In cases where a statute is relevant, many formalists believe that the most determinate mechanism for determining the content of a legal rule is by looking at the "plain language" of the statute.³⁰³ However, formalism and textualism are not coterminous. Formalist approaches need not be textualist. For example, in cases where the statute clearly adopts a standard (reasonableness, for example), the formalist has to go beyond the statute and develop bright-line common law rules. Indeed, as we will see, the Federal Circuit has been quite active in developing such rules.³⁰⁴

In several major areas of patent law decisionmaking—information technology, the intersection of patent and antitrust, and, most recently, the construction of the experimental use defense—the Federal Circuit's

303. The most prominent contemporary exponent of both textualism and rule-formalism is doubtless Supreme Court Justice Antonin Scalia. See Antonin Scalia, *Common-Law Courts in a Civil-Law System, in A Matter of Interpretation: Federal Courts and the Law* 3, 25 (Amy Gutmann ed., 1997) ("Of all the criticisms leveled against textualism, the most mindless is that it is 'formalistic.' The answer to that is, *of course it's formalistic!* The rule of law is *about* form. . . . Long live formalism. It is what makes a government a government of laws and not of men.") (emphasis in original). Formalism generally, and the textualist approach to statutory interpretation in particular, have been criticized by proponents of "dynamic" or "pragmatic" approaches to judicial decisionmaking. See, e.g., William N. Eskridge, Jr., Norms, Empiricism, and Canons in Statutory Interpretation, 66 U. Chi. L. Rev. 671 (1999) (arguing that subjecting theories of statutory interpretation to empirical, fact-based analysis might expand the role of pragmatist interpretation); William N. Eskridge, Jr., Textualism, The Unknown Ideal?, 96 Mich. L. Rev. 1509 (1998) (reviewing Scalia's *A Matter of Interpretation* and noting problems with the logic, determinacy, neutrality and legitimacy of Scalia's textualism); William N. Eskridge, Jr. & Philip P. Frickey, Statutory Interpretation as Practical Reasoning, 42 Stan. L. Rev. 321, 322 (1990) ("[F]oundationalism is a flawed strategy for theorizing about statutory interpretation. . . . [A] more modest approach, grounded upon 'practical reason,' is both more natural and more useful."). I discuss the arguments made by these scholars in *infra* Part III.B.1.

304. It is also conceivable that a textualist approach does not have to be formalist. A textualist might argue that policy-oriented court decisions that move beyond the statutory language are acceptable in cases where the legislature, by adopting a vaguely worded standard, has implicitly delegated policymaking power to the court. Indeed, this appears to be precisely how textualists have reconciled themselves to the non-rule based nature of certain fields of law, principally antitrust. See *infra* text accompanying notes 374–376.

The relationship between formalist and textualist decisionmaking might be approximated as follows:

approach to decisionmaking has been decidedly formalist. Consider the arguments made by the court in its prominent *State Street Bank & Trust Co. v. Signature Financial Group, Inc.* decision.³⁰⁵ In that case, the court rejected the idea, embodied in previous case law (including Supreme Court case law) that either mathematical algorithms or business methods pose any special concerns for purposes of the subject matter requirement of the patent statute. The subject matter requirement, which is codified in section 101 of the 1952 Patent Act, states: "Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor"³⁰⁶ The *State Street* court's analysis focused on what it alleged to be the "plain and unambiguous meaning of § 101."³⁰⁷ According to the court, this plain meaning showed that, so long as an invention could be considered a process, machine, manufacture, or composition of matter, section 101 placed no restrictions on its patentability.³⁰⁸

Unfortunately for the Federal Circuit's formalist conclusion, a variety of precedent—both Supreme Court precedent and precedent from the Federal Circuit's predecessor court, the Court of Customs and Patent Appeals—placed limits on the patentability of both algorithms and business methods. With respect to algorithms, this prior precedent suggested that algorithms were patentable only to the extent that they were embodied in a physical element, in the case of a product patent, or applied to a physical process, in the case of a process patent.³⁰⁹ In adopting the physicality

	Textualist	Non-Textualist
Formalist	Standard approach	Bright-line common law
Non-formalist	Policy-oriented common law where statutory language delegates policymaking responsibility to the courts	Policy-oriented common law

In this Article, I argue that the Federal Circuit should be engaging in policy-oriented common law development (lower right hand corner).

305. 149 F.3d 1368 (Fed. Cir. 1998).

306. Patent Act of 1952, Pub. L. No. 593 § 101, 66 Stat. 792, 797 (1952) (codified at 35 U.S.C. § 101 (2000)); *State Street*, 149 F.3d at 1372 (quoting Patent Act).

307. 149 F.3d at 1372.

308. Contrary to textualist tenets, the court did invoke legislative history to bolster its claim about the expansiveness of section 101. *Id.* at 1373 (citing Committee Reports that stated Congress intended statutory subject matter to "include anything under the sun that is made by man"). But it did this in the context of invoking a prior Supreme Court case that had cited this legislative history. *Id.* (invoking *Diamond v. Chakrabarty*, 447 U.S. 303, 309 (1980)).

309. In the relevant Supreme Court opinion, *Diamond v. Diehr*, the Court stated that where a process claim involves a mathematical algorithm, physical "[t]ransformation and reduction of an article to a different state or thing is the clue to . . . patentability." 450 U.S. 175, 184 (1981) (emphasis added, internal citations and quotation marks omitted). In *Diehr*, the physicality criterion was met because the patent claimed a process for curing rubber. The predecessor to the Federal Circuit, the Court of Customs and Patent Appeals (CCPA), had used the physicality criterion in cases involving not only process claims but also product claims. When a product claim invoked a mathematical algorithm, the CCPA

requirement, the Supreme Court had wanted to cabin the extent to which broad ideas that might serve as the foundation for much subsequent invention could be patentable.³¹⁰ In order to evade this precedent, the Federal Circuit mischaracterized the prior cases as standing merely for the proposition that patented inventions had to meet a very low-threshold—indeed nominal—“usefulness” requirement.³¹¹

To be sure, the “physicality” criterion was an imperfect mechanism for cabining the scope of patents on mathematical algorithms. As the Federal Circuit’s pre-*State Street* jurisprudence demonstrated, this requirement could be manipulated so as to allow, inter alia, the patentability of virtually any algorithm that was implemented on a computer.³¹² Moreover, under the patent statute, patent scope is usually restricted through section 103 nonobviousness requirements and section 112 disclosure requirements, not through the section 101 subject matter requirement.³¹³ But rather than explaining its departure from precedent, the *State Street* court incorrectly recharacterized the precedent as supporting its formalist conclusion. Indeed, nothing in the court’s opinion even acknowledged the step the court was taking in holding that all mathematical algorithms with any plausible “usefulness” are patentable subject matter.

Because *State Street* involved a claim to a *product* (specifically a machine that computed the daily value of mutual fund shares),³¹⁴ some commentators suggested that the Federal Circuit’s rejection of any physicality requirement with respect to *process* claims might be considered

held that the mathematical algorithm at issue had to be limited by a physical element. E.g., *In re Abele*, 684 F.2d 902, 906 (C.C.P.A. 1982).

310. *Diehr*, 450 U.S. at 185 (stating that “abstract ideas” are not patentable subject matter). In an earlier case, *Gottschalk v. Benson*, the Court had rejected as unpatentable subject matter an algorithm that converted decimal numbers into binary numbers. 409 U.S. 63 (1972). The Court emphasized the broad scope of patent, noting that such a patent might cover everything from “the operation of a train to verification of drivers’ licenses to researching the law books for precedents.” *Id.* at 68.

311. The court did not bother to explain how this usefulness inquiry might cohere with the utility requirement, which has long been considered a separate requirement for patentability. See *Rai*, *Patent Gold Rush*, *supra* note 5, at 201–11.

312. For example, in *Arrhythmia Research Technology, Inc. v. Corazonix Corp.*, the Federal Circuit held that a mathematical process for detecting and analyzing electrocardiographic signals was patentable subject matter. See 958 F.2d 1053 (Fed. Cir. 1992). According to the court, the physical transformation requirement could be satisfied by the fact that the process transformed an analog heart signal to a filtered digital signal. See *id.* at 1059. Under the reasoning of the *Arrhythmia* case, any algorithm that takes as its input data represented in electronic form and transforms the data in some way is patentable subject matter. In an earlier case, *In re Iwahashi*, the Federal Circuit had rejected an argument that the claimed product was an unpatentable mathematical algorithm by holding that the electronic circuitry necessary to perform the mathematical steps provided the necessary physical element limitation. See 888 F.3d 1370, 1375 (Fed. Cir. 1989).

313. See 35 U.S.C. §§ 103, 112, 101 (2000).

314. See *State St. Bank & Trust Co. v. Signature Fin. Group, Inc.*, 149 F.3d 1368, 1371–72 (Fed. Cir. 1998).

dicta.³¹⁵ Any room for speculation in that regard was eliminated the following year, when the Federal Circuit decided *AT&T Corp. v. Excel Communications, Inc.*³¹⁶ In *Excel*, the court held that a process claim involving a mathematical algorithm was patentable because it produced a useful result.³¹⁷ Unlike the opinion in *State Street*, the opinion in *Excel* did argue that rejecting the physicality requirement is not at odds with Supreme Court precedent, specifically the Court's 1981 decision in *Diamond v. Diehr*.³¹⁸ But the Federal Circuit's argument is simply incorrect. According to the Federal Circuit, *Diehr* merely suggested physical transformation as *one* mechanism for showing that a process involving a mathematical algorithm is patentable subject matter;³¹⁹ to the contrary, *Diehr* emphasized such physical transformation as "the clue to . . . patentability."³²⁰ Notably, the opinion in *Excel* squarely acknowledged the departure from CCPA precedent.³²¹ But the only significant attempt at explanation for this departure came at the conclusion of the opinion: There the court indicated that its rejection of prior limitations on the patentable subject matter requirement had the virtue of being a bright-line rule that should enable "reasonable prediction of outcomes."³²² In other words, the *Excel* court appeared eager to reject a standard—the physicality requirement—that had required factual inquiry into the nature of the invention at issue in favor of one that eliminated the need for such inquiry.

In embracing the patentability of business methods,³²³ the Federal Circuit has similarly opted for a bright-line rule over a more nuanced, case and fact specific approach. Indeed, the *State Street* opinion failed to discuss in any way the innovation policy implications of a decision to allow business method patents.³²⁴ The court's decision ignored the eco-

315. See *id.* at 1372 (noting that whether a machine or product is at issue "is of little relevance"); see also, e.g., Robert C. Haldiman, Prior User Rights for Business Method Patents, 20 St. Louis U. Pub. L. Rev. 245, 271 (2001) (stating that the *State Street* court, in dicta, did not find the difference between a machine and a process claim relevant).

316. 172 F.3d 1352 (Fed. Cir. 1999), cert. denied, 528 U.S. 946 (1999).

317. *Id.* at 1360–61.

318. See *id.* at 1356–60 (discussing *Diamond v. Diehr*, 450 U.S. 175 (1981)).

319. See *id.* at 1358–59.

320. See *Diehr*, 450 U.S. at 184 (emphasis added; quotation omitted).

321. See *Excel*, 172 F.3d at 1359.

322. *Id.* at 1361.

323. Notably, the *State Street* opinion denied that allowing business method patents represented any sort of legal change. The court argued that discussion of the business method exception in prior caselaw had been dicta. *State St. Bank & Trust Co. v. Signature Fin. Group, Inc.*, 149 F.3d 1368, 1375–76 (Fed. Cir. 1998). While some of this case language was perhaps dicta, other opinions specifically rejected business methods as patentable subject matter. See, e.g., *In re Sterling*, 70 F.2d 910, 911–12 (C.C.P.A. 1934) (denying protection to "ingenious and convenient arrangement" for transferring funds on the grounds that it lacked the requisite "physical structure").

324. See Leo J. Raskind, *The State Street Bank Decision: The Bad Business of Unlimited Patent Protection for Methods of Doing Business*, 10 Fordham Intell. Prop. Media & Ent. L.J. 61, 62 (1999) (noting that the *State Street* decision announces its "startling conclusion" without "explanation, limitation, [or] rationale").

nomic reality to which various commentators have pointed—namely, that patents are not generally necessary as a mechanism for promoting innovation in the area of business methods.³²⁵ In addition, as Rochelle Dreyfuss has noted, because business methods are often characterized by path dependence, the first mover advantage “goes a long way to assuring returns adequate to recoup costs and earn substantial profit.”³²⁶ A more nuanced, policy-oriented approach to business method patents might have looked to whether the software implementing the business method in question represented a nonobvious leap forward in terms of programming skill. To the extent that there was such a leap forward, a patent might be necessary to provide the requisite incentive. But such a policy-oriented approach to business method patents would have required inferior decisionmakers to engage in complex fact finding. Thus, absent competent fact finders, the Federal Circuit might be justified in refusing to adopt a more nuanced approach.

Patent law scholars may have gradually grown accustomed to the Federal Circuit’s formalism. Its decisions involving the intersection of patent and antitrust have, however, aroused the ire of antitrust scholars.³²⁷ While numerous authorities, including the Supreme Court and antitrust scholars across the ideological spectrum, have emphasized the

325. See, e.g., *id.* at 64 (“Both economic theory and empirical studies of patent-intensive industries cast doubt on the premise that patent protection of business methods is required as an incentive for innovation . . .”).

326. Rochelle Cooper Dreyfuss, *Are Business Method Patents Bad for Business?*, 16 *Santa Clara Computer & High Tech. L.J.* 263, 275 (2000). In the context of software-implemented business methods, for example, path dependence can arise because the first firm to popularize a particular business method will quickly develop an “installed base” of users who are unlikely to want to switch to a competitor. See, e.g., Jared Earl Grusd, *Internet Business Methods: What Role Does and Should Patent Law Play*, 4 *Va. J.L. & Tech.* 9, 43–49 (1999), at <http://www.vjolt.net/vol4/issues/v4i2a9-grusd.html> (on file with the *Columbia Law Review*) (noting the impact of the “head start advantage” on solidifying future customer base).

327. In a series of cases decided in the late 1990s, most notably *Nobelpharma AB v. Implant Innovations, Inc.*, the Federal Circuit decided that it would no longer apply the antitrust law of the circuit from which the case originated to antitrust questions raised by patent cases. 141 F.3d 1059, 1067 (Fed. Cir. 1998). Rather, in order to preserve uniformity, it would develop and apply its own antitrust law. The *Nobelpharma* decision overruled prior decisions in which the court had applied the law of the circuit in which the case arose to decide antitrust issues. See *Loctite Corp. v. Ultraseal Ltd.*, 781 F.2d 861, 875 (Fed. Cir. 1985) (“We must approach a federal antitrust claim as would a court of appeals in the circuit of the district court whose judgment we review.”). Some commentators have argued that the court’s expansion of its jurisdictional authority into antitrust is contrary to Congress’s intent when it created the Federal Circuit. See, e.g., James B. Gambrell, *The Evolving Interplay of Patent Rights and Antitrust Restraints in the Federal Circuit*, 9 *Tex. Intell. Prop. L.J.* 137, 139–44 (2001). Gambrell also points to the Federal Circuit’s decision to develop and use its own case law in several other areas at the perimeter of patent law. *Id.* at 144–48.

policy-oriented, common law nature of antitrust law development,³²⁸ the Federal Circuit has used a formalist approach.

For example, according to the Federal Circuit, the language of the patent statute, specifically § 271(d)(4), establishes that there are virtually no antitrust limitations on a patent owner's refusal to license or use any right to the patent.³²⁹ Consider the Federal Circuit's 2000 decision in *CSU, L.L.C. v. Xerox*.³³⁰ In that case, a group of independent service operators (ISOs) sued Xerox for refusing to sell or license to them replacement parts and diagnostic software for Xerox's high-speed copying machines and printers. The ISOs alleged that Xerox was violating section 2 of the Sherman Act in using its monopoly in the copier and printer area to monopolize or attempt to monopolize the service and parts market.³³¹ Xerox counterclaimed that the ISOs were violating its patents and copyrights on the parts and software. The Federal Circuit assumed jurisdiction of the appeal based on this counterclaim.³³²

In addressing the antitrust issue, the Federal Circuit invoked § 271(d)(4) in favor of Xerox's argument that because the parts and software were the subject of intellectual property (specifically patents and copyrights), antitrust law could place no limitations on Xerox's refusal to deal. The court noted that although Xerox's behavior might well have anticompetitive effect, it saw "no more reason to inquire into the subjective motivation of Xerox in refusing to sell or license its patented works than . . . in evaluating the subjective motivation of the patentee in bring-

328. See *State Oil Co. v. Khan*, 522 U.S. 3, 20–21 (1997) ("[T]he general presumption that legislative changes should be left to Congress has less force with respect to the Sherman Act in light of the accepted view that Congress 'expected the courts to give shape to the statute's broad mandate by drawing on common-law tradition.'" (quoting *Nat'l Soc'y of Prof'l Eng'rs v. United States*, 435 U.S. 679, 688 (1978))). Numerous antitrust scholars have also endorsed a common law approach to the Sherman Act. See, e.g., Frank H. Easterbrook, *Workable Antitrust Policy*, 84 Mich. L. Rev. 1696, 1705 (1986) ("The Sherman Act set up a common law system in antitrust."); John Shepard Wiley, Jr., *A Capture Theory of Antitrust Federalism*, 99 Harv. L. Rev. 713, 777 (1986) ("[The Sherman Act] is most reasonably interpreted to grant courts common law power to initiate substantive policy."); see also William N. Eskridge, Jr., *Dynamic Statutory Interpretation*, 135 U. Pa. L. Rev. 1479, 1516–17 (1987) (citing Judge Richard Posner's observation that antitrust law should be interpreted based on common law principles).

329. Section 271(d) states that "[n]o patent owner otherwise entitled to relief . . . shall be denied relief or deemed guilty of misuse or illegal extension of the patent right by reason of his having . . . refused to license or use any rights to the patent" 35 U.S.C. § 271(d) (2000).

330. 203 F.3d 1322 (Fed. Cir. 2000).

331. In *Eastman Kodak Co. v. Image Technical Servs., Inc.*, the Supreme Court had held that a single brand of a product could constitute a primary market and that the servicing of, and parts for, that brand could constitute a secondary market. 504 U.S. 451, 481–82 (1992).

332. Notably, as discussed further below, the recent Supreme Court case of *Holmes Group, Inc. v. Vornado Air Circulation Systems, Inc.* has eliminated the Federal Circuit's ability to assume jurisdiction based on a counterclaim. See 535 U.S. 826 (2002); *infra* note 393 and accompanying text.

ing suit to enforce that same right.”³³³ The only antitrust limitations on patentee behavior that the Federal Circuit has recognized are those that the Supreme Court has explicitly required lower courts to recognize—namely, that the patent holder cannot enforce its patent if fraud was involved in securing the patent, if the patentee is engaged in sham litigation, or if the patentee is using the patent to promote illegal tying.³³⁴ As Robert Pitofsky, former chairman of the Federal Trade Commission (FTC), has pointed out, the Federal Circuit’s rigid approach would probably excuse a host of other anticompetitive behavior that would be highly problematic under the antitrust law.³³⁵ For example, the Federal Circuit’s decision might condone a situation where a patent holder with market power refused to sell except on condition that the purchaser not buy from a competitor.³³⁶

By contrast, the Ninth Circuit approach presents an alternative doctrinal model. Although the Ninth Circuit has enunciated a strong presumption in favor of the patentee’s refusal to license being immune from antitrust scrutiny,³³⁷ this presumption can be overcome in cases where the evidence suggests that the patentee’s reliance on its patent right to justify the refusal to license is merely a pretext to mask anticompetitive conduct.³³⁸

Most recently, the Federal Circuit has displayed its formalist tendencies in its virtual elimination of the common law experimental use defense to patent infringement. As developed by Justice Story in the nineteenth century, the experimental use exemption was a narrow defense to infringement that allowed researchers who used a patented invention for purely noncommercial purposes to engage in such use without having to secure a license. When the Federal Circuit was first formed in 1982, this commercial/noncommercial distinction might have served as a relatively bright-line common law rule: Universities would be on one side of the divide and private firms on the other. As universities have increasingly been involved in commercial activity, however, what might have once been a bright-line rule has become difficult to implement without inquiring into the details of the research at issue. Uncomfortable with this state of affairs, the Federal Circuit recently moved to eliminate the exemption, at least for virtually all university research: In *Madey v. Duke University*, the Federal Circuit held that the experimental use defense can no longer be used by scientists or institutions whose mission it is to do research, no matter how noncommercial the research.³³⁹

333. *CSU*, 203 F.3d at 1327.

334. *Id.* at 1326–28.

335. See Robert Pitofsky, *Challenges of the New Economy: Issues at the Intersection of Antitrust and Intellectual Property*, 68 *Antitrust L.J.* 913, 921–23 (2001).

336. *Id.* at 922.

337. *Image Technical Servs., Inc. v. Eastman Kodak Co.*, 125 F.3d 1195, 1218 (9th Cir. 1997), cert. denied, 523 U.S. 1094 (1998).

338. *Id.* at 1219.

339. 307 F.3d 1351 (Fed. Cir. 2002).

That the Federal Circuit's jurisprudence would be narrowly formalist is not surprising. Tunnel vision is a well-known liability for specialist courts. By the same token, the *State Street*, *Xerox*, and *Madey* cases all expand quite significantly the domain of patent law. Moreover, the bar that practices before the Federal Circuit is strongly pro-patent.³⁴⁰ Thus one might argue that the Federal Circuit's decisionmaking reflects not so much simple formalism as a policy of high patent protection masquerading as formalism. In other words, the Federal Circuit has been captured by the patent bar and is using a thin veneer of formalism to disguise this capture. On this view, capture may have occurred because the patent bar was able to pressure the political branches into appointing pro-patent judges.³⁴¹ Alternatively, judges on the Federal Circuit may have started out agnostic but have become captured as a consequence of hearing disproportionately from the patent bar.³⁴² Because capture is an obvious alternate explanation for the court's jurisprudence, it is important to evaluate carefully, and consider the consequences of, this hypothesis.³⁴³

B. *Formalism or Capture?*

A number of empirical studies support the conclusion that the Federal Circuit tends to view patent holders favorably, at least in comparison with its predecessor generalist courts. For example, data collected by John Allison and Mark Lemley on district court and Federal Circuit deci-

340. Some have argued that because the patent bar obviously represents both patent holders and patent infringers, it should have no particular ideological disposition on patents. See, e.g., Dreyfuss, Federal Circuit, *supra* note 2, at 29–30. This view appears to be mistaken, however. The prevailing position in the patent bar appears to be that a strong patent system is essential for innovation. See, e.g., Edward G. Florio, Business Method Patents—Another Round of Patent Bashing 5 (Chair's Bulletin, ABA Section on Intellectual Property Law 2001) (decrying complaints that patent protection is too strong). As discussed earlier, this attitude may be a consequence of the fact that most alleged patent infringers are also patent holders. See *supra* text accompanying notes 176–177.

341. See, e.g., Richard A. Posner, *The Federal Courts: Challenge and Reform* 254–57 (1995) (arguing that because it is easier for lobbying groups to predict how judges will vote on a particular issue than on a range of issues, specialized courts are prone to capture). In the case of recent appointees to the Federal Circuit, however, it is not clear that the patent bar has had significant influence. See Harold C. Wegner, *Priority Patent Reform Initiatives* 6 (Dec. 2002) (unpublished manuscript, on file with the *Columbia Law Review*) (arguing that patent bar has not influenced selection of a Federal Circuit appointment since the appointment of Judge Alan Lourie in 1990).

342. A thoroughgoing public choice theorist might also argue that specialist patent judges interested in maximizing their own narrow self-interest—in this case, their power—would tend to have a maximalist vision of patent law. As discussed above, it is unlikely that narrow self-interest is the primary motive that guides decisionmakers. See *supra* notes 145–146 and accompanying text. Nonetheless, it would be unwise to rule out this motive altogether.

343. Indeed, it is the explanation to which various commentators have turned. See, e.g., Allan N. Litmann, *Restoring the Balance of Our Patent System*, 37 *IDEA* 545, 552–70 (1997) (describing how the Federal Circuit has upset doctrinal stability by favoring patent holders).

sions from 1989 to 1996 indicate that the overall rate of courts determining validity in a manner favorable to patentees is higher after the creation of the Federal Circuit than it was before.³⁴⁴ Moreover, a recent study by Glynn Lunney on the Federal Circuit's treatment of patent validity indicates that the percentage of patents held invalid, where validity was at issue and decided, decreased from fifty percent in the immediate pre-Federal Circuit era to twenty-five percent in 1994–1995.³⁴⁵ The role of the nonobviousness requirement in invalidating patents has declined particularly steeply. While obviousness was the basis for finding invalidity in the majority of appellate cases decided prior to the Federal Circuit era, it was a basis for invalidity in only twenty percent of the cases decided by the Federal Circuit in 1994–1995.³⁴⁶ Finally, a study conducted by Donald Dunner of Federal Circuit decisions from 1982 to 1994 noted that the court was “significantly more likely to affirm district court judgments in favor of patent owners than accused infringers.”³⁴⁷

One might also look to doctrinal trends in the Federal Circuit as evidence of capture. As noted earlier, the Federal Circuit has raised the evidentiary standard for proving a patent invalid to one that requires clear and convincing evidence of invalidity.³⁴⁸ In addition, the Federal Circuit has elevated the status of secondary considerations³⁴⁹ so that they play a role comparable to that played by the inquiry into the prior art.³⁵⁰ As various commentators have emphasized, reliance on secondary considerations, particularly commercial success, will tend to lead to conclusions of nonobviousness even in cases where the commercial success resulted from marketing or other factors unrelated to any technical advance represented by the invention. In all likelihood, only those patents that cover commercially successful technology will justify the expense of litigation.³⁵¹

These statistical and doctrinal shifts, coupled with scattered commentary from particular Federal Circuit judges, suggest that at least some members of the Federal Circuit view patent rights as a relatively unalloyed

344. See Allison & Lemley, *supra* note 56, at 205–06. This result is perhaps not surprising, given that the Federal Circuit has determined that issued patents should be given a very high presumption of validity. See *supra* notes 91–92 and accompanying text.

345. Glynn S. Lunney, Jr., *E-Obviousness*, 7 Mich. Telecomm. & Tech. L. Rev. 363, 371 & fig.1 (2001).

346. *Id.* at 373.

347. See Donald Dunner et al., *A Statistical Look at the Federal Circuit's Decisions: 1982–1984*, 5 Fed. Cir. B.J. 151, 154 (1995).

348. See *supra* notes 91–92 and accompanying text.

349. For a discussion of the role of secondary considerations in patent law, see *supra* note 59 and accompanying text.

350. See Lunney, *supra* note 345, at 375–78; Robert P. Merges, *Commercial Success and Patent Standards: Economic Perspectives on Innovation*, 76 Cal. L. Rev. 805, 823–27 (1988).

351. See Edmund W. Kitch, *Graham v. John Deere Co.*: New Standards for Patents, 1966 Sup. Ct. Rev. 293, 323–43 (discussing high costs of patent infringement suits).

good, comparable to rights in tangible property.³⁵² While conventional economic analysis of patents is concerned with the deadweight loss and impediments to future innovation that patents may create,³⁵³ and hence asserts that patents should issue only to the extent they are necessary to induce invention,³⁵⁴ an "ordinary property" view dismisses the possibility that patents create monopoly-like difficulties. Thus there is no reason to object to patents issuing for virtually all invention that is novel. The most theoretically sophisticated version of the "patents as ordinary property" view, propounded by Edmund Kitch, argues that patents should be granted early in the research and development process and should be broad, so as to promote future development and to allow the patent holder to coordinate this future development.³⁵⁵

Tendencies on the part of the Federal Circuit to view patents as ordinary property may be suggestive of capture; they do not, however, represent dispositive evidence of capture. In order to argue that these tendencies are dispositive evidence of capture, one would have to assume that no fair-minded court that had listened carefully to all perspectives on the question of how best to promote innovation could have reached an alternate conclusion. In addition, although there are reasons to be concerned about capture, the capture theory does not provide the most comprehensive explanation of the court's jurisprudence. For example, the court has clearly not accepted the most assertive version of patents-as-ordinary-property claim, which counsels in favor of broad grants. To the contrary, the Federal Circuit has on repeated occasions attempted to police broad claiming by patentees. This policing effort has been particularly marked in the area of biotechnology. The Federal Circuit's application of section 112 disclosure requirements, in particular the written description component of disclosure, appear to require biotechnology

352. See *Carl Schenck, A.G. v. Norton Corp.*, 713 F.2d 782, 786 n.3 (Fed. Cir. 1983) ("A patent, under the statute, is property Nowhere in any statute is a patent described as a monopoly. The patent right is but the right to exclude others, the very definition of 'property.'"); Lunney, *supra* note 345, at 367 & n.19 (citing speech by then-Chief Judge Markey in 1983). The *Carl Schenck* opinion was authored by Chief Judge Markey.

353. See, e.g., Subcomm. on Patents, Trademarks, and Copyrights of the Senate Comm. on the Judiciary, 85th Cong., *An Economic Review of the Patent System* 55, 64 (GPO 1958) (review written by Fritz Machlup) ("[E]xisting patents impose a burden on society, a burden which it has decided to carry in order to hold out to people the chance of obtaining future profits from future patents on future inventions."); F. M. Scherer, *Industrial Market Structure and Economic Performance* 450-53 (2d ed. 1980) (discussing social costs of the patent system, the most obvious being resource misallocation due to unnecessary patent protection).

354. Scherer, *supra* note 353, at 442 ("Ideally, the life of a patent should be no longer than it needs to be to encourage the optimal amount of invention, so that monopolistic restrictions are terminated as soon as possible.").

355. See Edmund W. Kitch, *supra* note 10, at 276. Kitch notes that he intends to "reintegrate[] the patent institution with the general theory of property rights." *Id.* at 265. Theories of tangible property rights also argue that such rights are given in order to induce socially useful investment in the property. E.g., Harold Demsetz, *Toward a Theory of Property Rights*, 57 *Am. Econ. Rev.* 347, 351-56 (1967).

patentees to disclose the precise structure of any invention that they are claiming.³⁵⁶ Claims that encompass more than the disclosed structure will be struck down as insufficiently described.

The Federal Circuit has also attempted to limit quite substantially scope-expanding doctrines such as the doctrine of equivalents. The doctrine of equivalents is a longstanding judge-made doctrine that prevents an accused infringer from evading infringement of a prior patent claim by making insubstantial changes that take her invention outside the claim's literal limits.³⁵⁷ The doctrine reflects a recognition of the limitations of language: In the words of the Supreme Court, "[t]he language in the patent claims may not capture every nuance of the invention or describe with complete precision the range of its novelty."³⁵⁸ As a consequence, if a patent's scope were limited to its literal claim terms, the value of the patent would be substantially diminished.³⁵⁹ Even while acknowledging the solid pedigree of the doctrine of equivalents, the Federal Circuit has tried to limit its force. For example, in the recent case of *Johnson & Johnson Associates v. R.E. Service Co.*, the Federal Circuit declared that a patent holder could not invoke the doctrine of equivalents to encompass subject matter it had previously disclosed but not claimed.³⁶⁰ The decision stands in tension with prior Supreme Court precedent suggesting that such subject matter could be encompassed under the doctrine of equivalents.³⁶¹ More dramatically, in its 2000 en banc decision *Festo Corp. v. Shoketsu Kinzoku Kogyo Kabushiki Co.*, the Federal Circuit majority upset many patent holders by establishing a bright-line bar against the assertion of the doctrine of equivalents in any circumstance where the relevant claim element had been amended in the patent prosecution.³⁶² As the en banc majority acknowledged, its decision overturned a line of cases in which the court had adopted a flexible bar that allowed some range of equivalents in certain circumstances.³⁶³ The majority argued, however, that the flexible bar had proven "unworkable" because it was "virtually impossible to predict before the decision on appeal where the line of surrender is drawn."³⁶⁴ To the extent that the Federal Circuit has been

356. See discussion *supra* notes 136–137, 165 and accompanying text.

357. See, e.g., *Graver Tank & Mfg. Co. v. Linde Air Prods. Co.*, 339 U.S. 605, 608–11 (1950).

358. *Festo Corp. v. Shoketsu Kinzoku Kogyo Kabushiki Co.*, 535 U.S. 722, 731 (2002).

359. *Id.*

360. 285 F.3d 1046, 1054–55 (Fed. Cir. 2002).

361. In *Graver Tank*, the Supreme Court allowed the patentee to use the doctrine of equivalents to encompass disclosed matter that had not been claimed. *Graver Tank*, 339 U.S. at 608–10.

362. 234 F.3d 558, 558 (Fed. Cir. 2000) (en banc). This decision was subsequently overturned by the Supreme Court. See 535 U.S. at 722; *infra* text accompanying notes 395–396.

363. *Festo*, 234 F.3d at 572–75.

364. *Id.* at 575. Notably, however, the majority provided no evidence for its conclusion that the district courts had failed to apply the flexible bar in a reasonably predictable fashion. The failure to present any evidence suggests that at least some Federal

attempting through decisions like *Festo* to establish bright-line rules that cabin the scope-expanding tendencies of equitable standards like the doctrine of equivalents, its jurisprudence has clearly been formalist rather than pro-patent.³⁶⁵

Even the Federal Circuit's increased reliance on secondary considerations³⁶⁶ may be considered formalism of a sort. To be sure, because secondary considerations are not mentioned in the patent statute, the court's formalism in this case is not explicitly textualist. Nonetheless, to the extent that secondary considerations are relatively bright-line, objective rules that do not require inquiry into complex facts (and the Federal Circuit clearly thinks that they are),³⁶⁷ the Federal Circuit's reliance on secondary considerations constitutes formalism.

At bottom, then, the Federal Circuit's major decisions in recent years may have been influenced by bias toward patent holders. Certainly the Federal Circuit's disproportionate exposure to a bar that is pro-patent should make us concerned about the court's pro-patent tendencies. Bias is, however, underinclusive as a general explanation. Moreover, those who would charge bias are vulnerable to the countercharge that they are smuggling into their institutional analysis a substantive position on appro-

Circuit judges may have a priori hostility to flexible standards. Indeed, one former Federal Circuit judge has called into question doctrine of equivalents jurisprudence more generally, on grounds that the standard undermines the "Rule of Law" within the patent system. See S. Jay Plager, Challenges for Intellectual Property Law in the Twenty-First Century: Indeterminacy and Other Problems, 2001 U. Ill. L. Rev. 69, 74, 81.

365. In a recent article, Polk Wagner argues that the Federal Circuit's bright-line approach in *Festo* can best be justified not as an attack on the doctrine of equivalents but, rather, as an attempt to police broad claiming by patentees. R. Polk Wagner, Reconsidering Estoppel: Patent Administration and the Failure of *Festo*, 151 U. Pa. L. Rev. 159, 165 (2002). In Wagner's view, if the penalty attached to claiming broadly and then having to amend one's claim in response to an examiner's objections is draconian, then the patentee will be wary of claiming broadly. Broad claiming will be chilled even if (as is likely, given the PTO's limited resources) the percentage of broad claims that are actually targeted for amendment is small. See *id.* at 216-17. The difficulties associated with policing behavior through low-probability, high-impact sanctions are well known, however. Such sanctions can distort primary behavior in undesirable ways. In the case of patent behavior, risk-seeking applicants may continue to claim broadly, while risk-averse applicants may claim too modestly. Wagner's larger point about the need to remedy the information asymmetry between the PTO and the patent applicant is well taken. See *id.* at 192-209. It would appear, however, that an opposition proceeding could remedy this asymmetry while causing fewer distortions in primary behavior. Cf. *supra* notes 186-191 and accompanying text (discussing opposition proceedings).

366. See *supra* text accompanying notes 60, 349-351 (discussing roles of secondary considerations in Federal Circuit jurisprudence).

367. The Federal Circuit now terms secondary considerations "objective evidence." See, e.g., *Hybritech, Inc. v. Monoclonal Antibodies, Inc.*, 802 F.2d 1367, 1380 (Fed. Cir. 1986) ("Objective evidence such as commercial success, failure of others, long-felt need, and unexpected results must be considered *before* a conclusion on obviousness is reached and is not merely 'icing on the cake.'"). This renaming points not only to the court's view of these considerations as establishing clear rules for patent validity but also to the court's elevation of these considerations from a secondary role.

priate levels of patent protection. By contrast, there can be no serious dispute that the Federal Circuit's jurisprudence is formalist in its orientation.

Unlike bias, formalism is of course eminently defensible as a normative matter. However, in the next section I argue that despite its considerable attractions, formalism is nonetheless a poor fit for the patent statute.

C. *The Limits of Formalism in Patent Law*

1. *Formalism Generally: Pro and Con.* — In general, formalism has many arguments in its favor. As an initial matter, to the extent that formalism employs textualist approaches to statutory interpretation, it can be defended on constitutional grounds. Textualists note that materials outside the language of the statutory text are an illegitimate source of law under Article 1, Section 7 of the Constitution.³⁶⁸ They also argue that requiring adherence to text addresses separation-of-powers and democratic legitimacy concerns by preventing judges from advancing their own policy preferences, as contrasted with those of the legislature.³⁶⁹ Proponents of rule-formalism also have functional arguments in their favor. Although formalism concededly does not provide the optimal resolution for particular cases, it can reduce judicial decisionmaking costs³⁷⁰ and increase predictability and stability in the law.³⁷¹ In addition, formalism can be defended on grounds of comparative institutional competence: Unlike legislatures, judges do not have the resources to engage in sound policymaking.

Critics of the textualist version of formalism argue that because the apparent "plain language" of statutes is often susceptible to different interpretations, such language does not constrain judicial discretion.³⁷² In addition, as a functional matter, because the determinacy of bright-line

368. See, e.g., Scalia, *supra* note 303, at 29–37 (discussing illegitimacy of using extratextual sources such as legislative history in interpreting statutes); see also William N. Eskridge, Jr., *The New Textualism*, 37 *UCLA L. Rev.* 621, 671–73 (1990) [hereinafter Eskridge, *New Textualism*] (discussing this argument).

369. Eskridge, *New Textualism*, *supra* note 368, at 673–75 (discussing this argument). For the classic statement of legitimacy concerns that arise from judges exercising policymaking power, see Alexander M. Bickel, *The Least Dangerous Branch: The Supreme Court at the Bar of Politics* (2d ed. 1986).

370. See Alexander, *supra* note 301, at 531, 544–45 (noting that although rule formalism will not always achieve the optimal resolution in a particular case, it does promote coordination and reduce the need for costly deliberation).

371. See Farber, *supra* note 301, at 542–43 (discussing formalist view that a system of judicial discretion "seems to provide little in the way of certainty, stability, or notice").

372. Cf. Eskridge & Frickey, *supra* note 303, at 341 (noting that even if language itself is not "intrinsically indeterminate," the types of politicized words often used in statutes are susceptible to different interpretations). This criticism of formalism dates back to the Legal Realists, who were the first to argue that, because statutory language is intrinsically indeterminate, judges can not avoid being policymakers. See, e.g., Karl Llewellyn, *The Bramble Bush* 76–79 (1930) (noting the judicial need to construct language in order to apply broadly applicable statutes, written *ex ante*, to a particular case).

rules (whether statute-based or common law) can be illusory, rules do not necessarily increase predictability and decisionmaking efficiency. Finally, although judges may not have the technical competence to be policymakers, judicial policymaking may be less subject to interest group rent-seeking than policymaking by political actors.³⁷³

In the abstract, the debate surrounding formalism may be difficult to resolve. For a variety of reasons, however, even those generally disposed to formalism might concede that patent law is not, and should not be, formalist in nature. In what follows, I argue that, for both structural and functional reasons, judicial decisionmaking in the area of patent law does have, and should have, a significant, self-conscious policy component. Indeed, although the structural argument has thus far been largely ignored by patent commentators, it is one that is very difficult to refute.

2. *Structural Arguments Against Patent Formalism.* — Even many committed formalists and textualists would probably agree that, in areas of law like antitrust, where Congress has delegated policymaking power to the courts, separation of powers arguments in favor of textualism have diminished weight. In the words of William Baxter, an antitrust official in the Reagan Administration, the antitrust statutes are a type of “enabling legislation that has permitted a common-law refinement . . . through an evolution guided by only the most general statutory directions.”³⁷⁴ In short, any separation of powers principles that might counsel against judicial policymaking have long since given way in the context of antitrust. Thus, in cases that involve the intersection of patent and antitrust (as many do, given that antitrust law is the central legal doctrine limiting patent scope), even formalists should accept that at least some common law development is legitimate.

More generally, most scholars agree that the creation of federal common law may be legitimate where a specific federal enactment can be interpreted as authorizing its creation.³⁷⁵ This principle has relevance even in patent cases where antitrust issues are not raised explicitly. This is because the legislative history and language of certain patent law provisions, particularly provisions governing patent validity and scope, strongly indicate that Congress wanted the courts to engage in relatively wide-ranging interpretation of these provisions. To be sure, patent law is not

373. See *supra* text accompanying notes 141–148.

374. William F. Baxter, *Separation of Powers, Prosecutorial Discretion, and the “Common Law” Nature of Antitrust Law*, 60 *Tex. L. Rev.* 661, 663 (1982); see also sources cited *supra* note 303.

375. See, e.g., Curtis A. Bradley & Jack L. Goldsmith, *Customary International Law as Federal Common Law: A Critique of the Modern Position*, 110 *Harv. L. Rev.* 815, 856 (1997) (“Courts and scholars generally agree that federal common law must be authorized in some fashion by the Constitution or a federal statute.”); Martha A. Field, *Sources of Law: The Scope of Federal Common Law*, 99 *Harv. L. Rev.* 881, 887 (1986) (noting that the consensus among modern scholars is that, “to create federal common law, [a federal court] must point to a federal enactment, constitutional or statutory, that it interprets as authorizing the federal common law rule”).

precisely equivalent to antitrust law. We may still be in the realm of legal interpretation, and not in the realm of the wide-ranging common law creation that arguably distinguishes antitrust law. Nonetheless, the difference is one of degree—and a relatively small degree at that—not kind.³⁷⁶

Consider, for example, the critical validity requirement of nonobviousness. The heart of the nonobviousness requirement is contained in two sentences that read as follows:

A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.³⁷⁷

The spareness of the statutory language suggests the need for further judicial elaboration. Indeed, at the time the 1952 patent statute was enacted, the eminent patent lawyer P. J. Federico wrote an accompanying commentary, considered part of the legislative history of the 1952 Act, that made precisely this point. In this commentary, Federico noted that “[t]he problem of what is obvious and hence not patentable is still of necessity one of judgment. The statute does not purport to categorize the particular criteria according to which the [nonobviousness] judgment is to be exercised. . . .”³⁷⁸ Moreover, the statutory nonobviousness requirement is itself a creature of the common law. Prior to the requirement’s inclusion in the 1952 Patent Act, some type of nonobviousness-type language had been used in judicial opinions for over a hundred years.³⁷⁹ As the House and Senate Reports on the 1952 Patent Act point out, the statutory language merely “paraphrases language which has often been used in decisions of the courts”³⁸⁰ To the extent that Congress

376. See Peter Westen & Jeffrey S. Lehman, *Is There Life for Erie After the Death of Diversity?*, 78 Mich. L. Rev. 311, 339–41 (1980) (noting that statutory interpretation and common law creation lie on a spectrum).

377. 35 U.S.C. § 103(a) (2000).

378. P.J. Federico, *Commentary on the New Patent Act*, 75 J. Pat. & Trademark Off. Soc’y 161, 184 (1993). More generally, Federico’s commentary notes that “[w]hile patents are creatures of statute, the entire body of patent law is much fuller than the statute itself, including a vast amount of case material on subjects such as invention and infringement which are dealt with by the statute only in general terms.” *Id.* at 162. The commentary is considered to have authority at least equal to that of the somewhat sparse House and Senate reports on the 1952 Act. See Staff of the Senate Comm. on the Judiciary, 82d Cong., Revision of Title 35, United States Code: Report to Accompany H.R. 7794 (Comm. Print 1952) [hereinafter Senate Report]; Staff of the House Comm. on the Judiciary, 82d Cong., Revision of Title 35, United States Code “Patents”: Report to Accompany H.R. 7794 (Comm. Print 1952) [hereinafter House Report].

379. See Senate Report, *supra* note 378, at 6; Federico, *supra* note 378, at 181–83 (discussing 19th century Supreme Court cases suggesting a nonobviousness requirement).

380. Senate Report, *supra* note 378, at 6; House Report, *supra* note 378, at 7.

indicated interest in setting its own substantive standard for nonobviousness, it was only to indicate that an invention did not (as a few court opinions had suggested) have to involve a "flash of genius" in order to be patentable.³⁸¹ Since 1952, the nonobviousness standard has been developed further by various courts, including the Supreme Court.³⁸² By contrast, Congress has since shown little interest in specifying the standard further. The common law origins of the nonobviousness requirement, and Congress's lack of interest in specifying the content of the requirement as a statutory matter, strongly suggest that it would be appropriate for judges to interpret this requirement in a policy-oriented manner.

Like the inquiry into nonobviousness, the inquiry into adequate disclosure is central to both patent validity and scope. The principal requirement of adequate disclosure is enablement. To the extent that the enablement requirement is high, only narrow patent claims will be held valid. Like the language of the nonobviousness requirement, the language of the enablement requirement is quite general, requiring only that the patentee show one of ordinary skill in the art how to "make and use [the invention]."³⁸³ Like nonobviousness, then, the enablement standard might be used by the courts in a policy-oriented manner, particularly for the purposes of determining patent scope. As economists and legal scholars across the ideological spectrum have noted, questions of patent scope are critical for purposes of regulating follow-on innovation.³⁸⁴

Indeed, in the area of patent scope, there is a long history of judge-made law, and congressional acceptance thereof.³⁸⁵ As an initial matter, the canons by which judges interpret patent claims are all judge-made law. Nothing in the patent statute suggests, for example, that a patent claim should be interpreted in light of the embodiments in the specification but that the claim is nonetheless not limited to those specific embodiments.³⁸⁶ To the contrary, these canons suggest a policy-driven attempt by judges to balance the need for certainty in claim scope with the need to give the innovator sufficient incentives. Another important determinant of claim scope, the doctrine of equivalents, is entirely a judicial crea-

381. House Report, *supra* note 378, at 18. The second of the two sentences of the nonobviousness requirement, which states, somewhat inartfully, that "patentability . . . is not to be negated by the manner in which the invention is made" speaks to the concern regarding genius. 35 U.S.C. § 103(a).

382. See, e.g., *Graham v. John Deere Co.*, 383 U.S. 1, 12-19 (1966). The Federal Circuit has also developed the nonobviousness requirement to some extent. In the case of the Federal Circuit, however, development has been constrained by its emphasis on bright-line rules. See *supra* text accompanying notes 156-159, 174-175, 366-367.

383. 35 U.S.C. § 112.

384. For discussion of the importance of patent scope, see *supra* text accompanying notes 174, 353-355.

385. See *Merges & Nelson*, *supra* note 164, at 841 (noting that the legal principles that govern patent scope "leave considerable room for discretion").

386. See *supra* note 43.

tion. As noted earlier, the innovation policy motivation behind this doctrine is clear: Given the inevitable limitations of language, there may be circumstances where a clever infringer can escape literal infringement of a patent claim but nonetheless essentially be doing what the patentee has done.³⁸⁷ If we were to allow the infringer to escape infringement in such situations, it would substantially diminish the incentives of the initial innovator. Moreover, although the Federal Circuit's formalist impulses have led it to attempt to cabin the scope of the doctrine of equivalents, Congress itself has shown no such inclination.

Finally, even sections of the statute that may appear relatively "bright line" in character are, upon closer examination, not so clear. Consider, for example, the subject matter requirement. According to the Federal Circuit, the statutory language regarding subject matter—which refers to "process[es], machine[s], manufacture[s], or composition[s] of matter"—can admit of only one interpretation. Moreover, this interpretation does not, for all practical purposes, place any substantive limits on patentable subject matter. As noted earlier, however, the subject matter requirement has historically not been interpreted in this manner. To the contrary, the Supreme Court has long interpreted the requirement to exclude broad, abstract ideas (including algorithms without a physicality limitation).³⁸⁸ In addition, in the recent case of *J.E.M. AG Supply, Inc. v. Pioneer Hi-Bred International, Inc.*, which involved the patentability of plants, the Supreme Court emphasized the dynamic and evolving nature of the subject matter provision.³⁸⁹ In the face of this precedent, the Federal Circuit's insistence on plain language as leading to a single conclusion regarding subject matter is misplaced.

There should be little question that the patent statute, as currently structured, contemplates policy-oriented judicial development of patent common law. Thus the formalist who is predominantly concerned with unauthorized judicial policymaking should not object to patent common

387. See *supra* notes 359–360 and accompanying text.

388. As discussed earlier, see *supra* note 310, the Court alluded to such concerns in *Gottschalk v. Benson*, 409 U.S. 63 (1972). The Court has also raised this type of argument in cases involving the utility requirement of the patent statute. For example, in *Brenner v. Manson*, the Court raised the concern that allowing patents on basic research that was primarily useful as the foundation for future research would "confer power to block off whole areas of scientific development, without compensating benefit to the public." 383 U.S. 519, 534 (1966) (citation omitted). It could be argued that the Supreme Court took a somewhat formal approach to the subject matter requirement in *Diamond v. Chakrabarty*, 447 U.S. 303 (1980). In that case, the Court rejected arguments that it should consider the policy consequences of genetic engineering research in determining whether such research should be patentable. The Court argued that such policy determinations should be left to Congress. *Id.* at 317–18. In *Chakrabarty*, however, the policy arguments in question were moral arguments wholly unrelated to the innovation goals of patent law. Although courts have been delegated the power to make common law that furthers the innovation goals of the patent statute, they have not been delegated common law power to make judgments about the morality of scientific research.

389. 534 U.S. 124, 135 (2001).

law. The fact that the patent statute appears to authorize policy-oriented common law development does not, however, answer the normative question of whether such development is a good idea.

3. *Functional Arguments Against Patent Formalism.* — In some respects, the Federal Circuit's formalist decisions themselves stand as the best arguments against a formalist approach. Each of its major formalist opinions has not only been criticized by numerous commentators but has also been called into question by other institutions within the patent system. For example, concern about the Federal Circuit's *State Street* decision on business method patents³⁹⁰ extends beyond the academic community: In particular, the Patent and Trademark Office's decision to put into place a special layer of additional review before granting business methods patents underscores the anomalous nature of such patents.³⁹¹

Similarly, the Federal Circuit's *CSU v. Xerox* decision, which declares a de facto patent law exception to antitrust principles in the context of refusals to deal,³⁹² has recently been significantly undermined by the Supreme Court. In *Holmes Group, Inc. v. Vornado Air Circulation Systems, Inc.*, the Supreme Court barred Federal Circuit jurisdiction over future cases like *Xerox*, where the patent issue arises only as a counterclaim.³⁹³ In his concurring opinion in that case, Justice Stevens emphasized that

the plaintiff's choice of forum includes not only the court that will conduct the trial but the appellate court as well. A plaintiff who has a legitimate interest in litigating in a circuit whose precedents support its theory of the case might omit a patent claim in order to avoid review in the Federal Circuit.³⁹⁴

Just as the Supreme Court has questioned the Federal Circuit's formalism in the statutory context, so too has it questioned this formalism in the common law context. The Federal Circuit's *Festo* decision, which sets up a bright-line bar against the assertion of the doctrine of equivalents in cases of claim term amendment has been overturned unanimously by the Supreme Court.³⁹⁵ In place of the bright-line bar, the Supreme Court has reinstated a more flexible standard that requires inquiry into particularized facts. This standard allows the assertion of the doctrine of

390. See *State St. Bank & Trust Co. v. Signature Fin. Group Inc.*, 149 F.3d 1368 (Fed. Cir. 1998), discussed *supra*, text accompanying notes 305–308.

391. See United States Patent and Trademark Office, USPTO White Paper: Automated Financial or Management Data Processing Methods (Business Methods), available at <http://www.uspto.gov/web/menu/busmethp/quality.htm#quality> (last visited May 7, 2003) (on file with the *Columbia Law Review*); Russel A. Korn, Comment: Is Legislation the Answer? An Analysis of the Proposed Legislation for Business Method Patents, 29 Fla. St. U. L. Rev. 1367, 1373 (2002).

392. See *CSU, L.L.C. v. Xerox*, 203 F.3d 1322 (Fed. Cir. 2000), discussed *supra* text accompanying notes 330–334.

393. 535 U.S. 826 (2002).

394. *Id.* at 1897 (Stevens, J., concurring).

395. See *Festo Corp. v. Shoketsu Kinzoku Kogyo Kabushiki Co.*, 234 F.3d 558, 574–78 (Fed. Cir. 2000) (en banc), vacated by 535 U.S. 722 (2002).

equivalents in those circumstances where “at the time of the amendment one skilled in the art could not reasonably be expected to have drafted a claim that would have literally encompassed the alleged equivalent.”³⁹⁶ In overturning the Federal Circuit decision, the Supreme Court argued that the same policy concerns that allow the assertion of the doctrine of equivalents in the ordinary case should also allow such an assertion when a claim limitation has been amended.

Of course, the fact that other institutions have questioned the Federal Circuit’s formalism does not render it invalid. The formalist might argue that although the Federal Circuit’s predilection for clear rules does not produce optimal results in particular cases, it reduces decisionmaking costs and increases predictability. Predictability is a particularly important value in the context of patent rights, which are supposed to provide a hedge against the risks inherent in the innovative process.³⁹⁷ Moreover, the failure to assign rights optimally can be mitigated by private sector rearrangement of such rights.

Even here, the formalist argument is suspect. Consider first the issue of decisionmaking costs. In the context of business method patents, the Federal Circuit’s formalist approach clearly limited the court’s own decisionmaking costs. By invoking formalism, the court did not have to engage in any inquiries about the economic issues surrounding such patents. In addition, inferior decisionmakers such as the lower courts and the PTO will be constrained in their decisionmaking on patentable subject matter. By the same token, however, the *State Street* decision has increased decisionmaking costs in the long term: It has engendered a flood of patent applications that will clog the PTO and court system for years to come.³⁹⁸ As for private sector rearrangement of assigned rights, such rearrangement clearly does not work under the conventional view of patents: If the purpose of patents is to provide incentives to invent, giving patents to inventions that would have been produced anyway entails unnecessary deadweight loss. (Alternatively, patents that are unduly narrow may fail to provide invention.) A reliance on private sector rearrangement of rights is more in keeping with the views of theorists like Edmund Kitch, who see patents as similar to ordinary property rights.³⁹⁹ But even Kitch appears to recognize that the transaction costs of reassignment are such that the scope of the initial allocation does matter (and should be broad).⁴⁰⁰

396. *Festo*, 535 U.S. at 741.

397. See supra note 167 and accompanying text.

398. John R. Allison & Emerson H. Tiller, Internet Business Method Patents, Tex. Bus. Rev. 1, Oct. 2002, at 1, available at http://www.utexas.edu/depts/bbr/tbr/Oct_02.pdf (on file with the *Columbia Law Review*) (noting that the number of business method patent applications has risen from 2,700 in 1999 to an expected 12,000 in 2001).

399. See generally Kitch, supra note 10 (discussing theory of patents as ordinary property).

400. *Id.* at 285. Kitch does argue that parties should be *permitted* to rearrange control of patents in order to achieve ownership of a broad prospect by one entity. But he does

In any event, bright-line rules are not the only mechanism for fostering predictability and certainty. To the contrary, the type of carefully formulated rebuttable presumption adopted by the Supreme Court in its *Festo* decision—barring assertion of the doctrine of equivalents in cases of amendment unless the patentee could not reasonably have drafted a claim that literally encompassed the alleged equivalent—can provide a fair amount of predictability without wholly sacrificing flexibility.⁴⁰¹ In fact, because patent law has for so long been a system of policy-oriented common law, the introduction of bright-line rules itself undermines predictability. In overturning the Federal Circuit's decision in *Festo*, the Supreme Court emphasized that the lower court's decision to impose a bright-line bar was a "[f]undamental alteration" that destroyed "the legitimate expectations of inventors in their property."⁴⁰² In this regard, it bears mention that a 1999 report by the Commission on Structural Alternatives for the Federal Courts of Appeals determined that the Federal Circuit was second only to the Ninth Circuit in terms of the proportion of attorneys who concluded that the law of the circuit was difficult to discern due to conflicting precedents.⁴⁰³

Finally, as noted earlier, to the extent that the fact-finding capacity of the inferior decisionmakers has been fortified, it is difficult to justify patent formalism on the grounds that these decisionmakers need to be tightly controlled. Even if, as is likely, application of the typical policy directive requires fact finding, the PTO and trial court reforms suggested above should give these inferior institutions the ability to find facts accurately.⁴⁰⁴

In sum, a substantial number of functional arguments mitigate against formalism in the arena of patent law.

D. Institutional Structure and Resources: The Comparative Analysis

The most compelling functional argument in favor of a formalist approach involves institutional structure and resources. The institutional limitations of courts in designing policy are well known.⁴⁰⁵ Courts can

not argue that such rearrangement will necessarily occur. See *id.* ("Ownership of different parts of what can be most efficiently exploited as one prospect may be in different hands. The only way to obtain the efficiency gains of a prospect may be to permit the parties to rearrange control of the various patents involved.").

401. See Phillip Areeda, *Antitrust Law as Industrial Policy: Should Judges and Juries Make It?*, in *Antitrust, Innovation, and Competitiveness* 29, 42 (Thomas M. Jorde & David J. Teece eds., 1992) ("Carefully formulated presumptions . . . can approach the clarity of bright-line rules without sacrificing flexible responsiveness to the peculiarities of a particular case.").

402. *Festo Corp. v. Shoketsu Kinzoku Kogyo Kabushiki Co.*, 535 U.S. 722, 739 (2002).

403. See Carl Tobias, *The White Commission and the Federal Circuit*, 10 *Cornell J.L. & Pub. Pol'y* 45, 58 (2000) (citing working papers of Commission).

404. See *supra* Part II.

405. These observations on the functional limitations of courts date back to the traditional model of litigation as a "vehicle for settling disputes between private parties

shape policy only through the mechanism of cases that actually come before them. These cases may not be representative of the range of issues that need to be decided. Relatedly, the only views that courts will generally hear are from the litigating parties. The litigating parties may not have an incentive to present all of the relevant arguments. Finally, and perhaps most importantly, the court system does not have the resources to evaluate fully complicated questions of economic policy. Indeed, this last limitation is particularly acute for appellate courts, which do not generally rely on court-appointed experts, special masters, or other mechanisms for sifting through complex technical information.

These arguments against judicial evolution apply with some force in the patent context. As noted earlier, the Federal Circuit is closely connected with the patent bar.⁴⁰⁶ Thus there may be an institutional bias towards patents. In addition, the Federal Circuit appears to be quite resistant to economic policy analysis. This resistance might well persist even if inferior decisionmakers were reformed. Indeed, even if patent formalism can only be *justified* by reference to questionable inferior decisionmakers, the Federal Circuit may well be committed to formalism for other reasons.

The Federal Circuit is not, however, the only possible vehicle for appellate review. Even within the current system, we have generalist review through the Supreme Court. Moreover, one might imagine a greater level of generalist review through one of the following three options: abolition of the Federal Circuit, such that appeals would go from the specialized trial court proposed in Part II to generalist circuit courts; the insertion of an additional layer of appellate review between the Federal Circuit and the Supreme Court; and an acceleration of the current trend towards more aggressive review of the Federal Circuit by the Supreme Court.

The most radical option for securing such generalist input would be abolition of the Federal Circuit. In other words, rather than having generalist trial courts reviewed by a specialized appellate court, we would have a specialized trial court reviewed by generalist appeals courts. Abolition of the Federal Circuit is an idea with considerable appeal. Because generalist courts regularly hear many cases, including antitrust cases, that require common law analysis, they may be more versed in such analysis than is the Federal Circuit. To be sure, in any given case, even a general-

about private rights." Abram Chayes, *The Role of the Judge in Public Law Litigation*, 89 Harv. L. Rev. 1281, 1282-83 (1976). On the traditional view, litigation is bipolar, retrospective, self-contained, party-initiated and party-controlled. *Id.* This traditional model of litigation, which tends to counsel in favor of formalist decisionmaking, came under attack as early as the turn of the twentieth century, with the ascendancy of the legal realists. See *supra* note 372. Moreover, as Chayes points out, various procedural changes, which allow courts to act, at least on occasion, in a quasi-legislative capacity, undermined the traditional model. Chayes, *supra*, at 1298-1304.

406. See *supra* text accompanying notes 176-177.

ist court might not hear legal and policy arguments against strong patent protection: If the parties to a given case were both patent holders, they might not be interested in presenting such arguments. This problem could, however, be remedied through such mechanisms as the filing of amicus briefs—particularly amicus briefs by agencies with economic expertise, such as the Federal Trade Commission or the Department of Justice's Antitrust Division.⁴⁰⁷ Moreover, because generalist courts would not have an intimate connection with the patent bar, they might be more likely than the Federal Circuit to pay attention to these amicus briefs.

Nonetheless, at the end of the day, abolishing the Federal Circuit raises difficulties. Most importantly, such abolition is probably a political nonstarter, given the significant influence of patent holding groups in Congress and the perception that the Federal Circuit is more likely to issue favorable decisions to patent holders.⁴⁰⁸ In addition, the transition problems associated with abolition might be significant: Now that we *have* the Federal Circuit, it may be too disruptive to do away with it. Finally, given the Supreme Court's historical lack of interest in taking significant numbers of patent cases, the Federal Circuit may be necessary in order to ensure uniformity in the basic legal principles that undergird patent law. There is some evidence (although the evidence is surprisingly quite limited) suggesting that, prior to the creation of the Federal Circuit, divergent patent law in different circuits spurred forum shopping and uncertainty; this evidence was important in the Congressional decision to create the Federal Circuit.⁴⁰⁹

A less radical mechanism for promoting evenhanded policymaking at the appellate level might involve regular en banc review of Federal Circuit decisions by a panel of judges drawn from circuits other than the Federal Circuit. In a recent essay, Michael Abramowicz offers this type of en banc review as a mechanism for reducing the impact of ideology on decisionmaking by appellate judges.⁴¹⁰ As he points out, existing en banc

407. See *infra* note 420 and accompanying text.

408. See *supra* text accompanying notes 340–355. In contrast, it should be politically possible to implement an opposition proceeding within the PTO or to establish specialized trial courts. As discussed earlier, even groups interested in high levels of patent protection should have some sympathy for accurate fact finding in specific cases, particularly if this fact finding can defeat their adversaries' patents. See *supra* text accompanying notes 186–187.

409. As Cecil Quillen has pointed out, the evidence of divergent decisionmaking presented to Congress was not statistically significant: Generalizations about the behavior of certain appellate courts were made on the basis of a handful of cases. Cecil D. Quillen, Jr., *The U.S. Patent System: Is it Broke? And Who Can Fix It if It Is?*, Address before the Spring Meeting of the Association of General Counsel 17–19 (May 11, 2001), available at <http://www.ftc.gov/os/comments/intelpropertycomments/quillenattachments/isitbrokewhocanfixit.pdf> (on file with the *Columbia Law Review*). Moreover, it does not appear that the Federal Circuit has done a particularly good job of establishing clarity in the patent law. See *supra* text accompanying notes 401–403.

410. Michael Abramowicz, *En Banc Revisited*, 100 *Colum. L. Rev.* 1600, 1616–17 (2000).

review with courts of appeal may be “too embedded in the politics of individual circuits to be an effective mechanism of discipline.”⁴¹¹ Abramowicz also notes that en banc review by outsiders might be particularly useful in providing for generalist review of specialized courts like the Federal Circuit.⁴¹² En banc review could play a role not only in defusing ideological content but also in tempering the Federal Circuit’s formalist inclinations. The prospect of en banc review by outsiders who were attentive to the innovation policy underpinnings of the statute would presumably spur the court to write opinions that were persuasive from a policy standpoint.⁴¹³

To the extent that such en banc review could, as a political matter, be implemented, such review would be a useful option for securing generalist input. Even without additional legislation, however, we could, and should, have aggressive review by the Supreme Court. Thus we should welcome the Supreme Court’s recent interest in reviewing patent cases that raise not only allocation of power issues but also issues of substantive patent law and policy. The Supreme Court’s recent interest in patent law is particularly welcome to the extent that at least some Justices on the Court have indicated an explicit awareness of the need to discipline the Federal Circuit. In the recent *Holmes v. Vornado* case, for example, Justice Stevens responded to the charge that the Court’s decision would divest the Federal Circuit of exclusive jurisdiction over all cases raising patent law issues by noting that other circuits *should* play a role in the development of patent law.⁴¹⁴ He observed that “[a]n occasional conflict in decisions may be useful in identifying questions that merit this Court’s attention. Moreover, occasional decisions by courts with broader jurisdiction will provide an antidote to the risk that the specialized court may develop an institutional bias.”⁴¹⁵ Thus, even if other generalist review is not forthcoming, the Supreme Court is in a position to shape Federal Circuit jurisprudence.

At bottom, then, the Federal Circuit does not have to be the circuit it currently is. Although the court may never be able to transcend entirely the constraints associated with specialization, the prospect of generalist en banc review, coupled with more rigorous Supreme Court review,

411. *Id.* at 1617.

412. *Id.* at 1624–25. The specifics of the Abramowicz proposal are not essential for our purposes here.

413. As the Hruska Commission, which studied the possibility of a specialized appellate court for patents in the 1970s, noted, “[v]esting exclusive jurisdiction over a class of cases in one court might reduce the incentive, now fostered by the possibility that another court will pass on the same issue, to produce a thorough and persuasive opinion.” Comm’n on Revision of the Federal Court Appellate System, Structure and Internal Procedures: Recommendations for Change, 67 F.R.D. 195, 234–36 (1975).

414. *Holmes Group, Inc. v. Vornado Air Circulation Sys., Inc.*, 535 U.S. 826, 838–39 (2002) (Stevens, J., concurring in part and concurring in the judgment).

415. *Id.* at 1898.

should dislodge the Federal Circuit from its rigid adherence to formalism.

Although greater generalist input would address the problems of limited vision and possible bias, such input would not address the fact that courts are systematically challenged with respect to the resources they have to evaluate complicated economic issues.⁴¹⁶ This is particularly true in the area of innovation, where the relevant economic issues in the various industries covered by the patent system can change quickly.⁴¹⁷ On the other hand, there is an enormous body of empirical and theoretical literature on the economic impact of patents.⁴¹⁸ Appellate courts could not only avail themselves of such literature but could also benefit from amicus briefs that explain this economic analysis.⁴¹⁹ In this regard, the recent interest in patent system reform of the FTC and the Antitrust Division of the Department of Justice (DOJ)⁴²⁰ is quite salutary. Both the FTC and the DOJ have strong traditions of economic analysis—traditions that are, unfortunately, as explained further below, absent at the Patent and Trademark Office itself.

One could imagine a variety of situations in which amicus briefs by such agencies like the FTC and the Antitrust Division of the DOJ could offer appellate courts assistance on economic analysis. For example, suppose a biopharmaceutical firm or university⁴²¹ had a patent on an invention that could plausibly be considered a research platform—the example of a patent on the NF- κ B cell signaling pathway held by Harvard, MIT, and the Whitehead Institute comes to mind.⁴²² The relevant patent claims all involve methods of blocking or reducing the activity of the NF- κ B pathway. Because this pathway has been implicated in an enormous variety of diseases, ranging from cancer and rheumatoid arthritis to osteoporosis and sepsis, the patent's claims encompass a huge range of potential therapies. Indeed, the patentees, together with their exclusive licen-

416. As the Supreme Court has noted in the context of antitrust law, "courts are of limited utility in examining difficult economic problems." *United States v. Topco Assocs., Inc.*, 405 U.S. 596, 609 (1972).

417. See Richard A. Posner, *Antitrust in the New Economy*, 68 *Antitrust L.J.* 925, 939 (2001) (noting the "mismatch between law time and new-economy real time").

418. See Craig Allen Nard, *Toward a Cautious Obeisance: The Role of Scholarship in Federal Circuit Patent Law Jurisprudence*, 39 *Houston L. Rev.* 667, 669–72 (2002) (listing scholarship).

419. See *id.* at 681 (noting that Second and Ninth Circuits rely on scholarship in their copyright and trademark opinions "considerably more often than the Federal Circuit has in its patent law opinions").

420. See Federal Trade Comm'n Hearings on Competition and Intellectual Property Law and Policy in the Knowledge-Based Economy (Feb. 6, 2002), available at www.ftc.gov/opp/intellect/index.htm (on file with the *Columbia Law Review*).

421. Universities have become increasingly active in patenting upstream research. See Arti K. Rai & Rebecca S. Eisenberg, *Bayh-Dole Reform and the Progress of Biomedicine*, 91 *Am. Scientist* 52, 52 (2003) (describing how Bayh-Dole Act turned universities into major players in the biotech business).

422. U.S. Patent No. 6,410,516 (issued June 25, 2002).

see Ariad Pharmaceuticals, are currently suing Eli Lilly for infringement and have sent letters alleging infringement to dozens of other companies.⁴²³ In an infringement case, the question of whether the claims of the NF- κ B patent were valid would be unclear: The Federal Circuit has not decided whether a patentee can claim all methods of inhibiting a pathway merely by identifying the pathway and its role in various diseases. The appellate court might, in considering issues of patent scope, give some weight to the considerable economic literature that argues against affording broad scope to patents that represent a research platform, on the theory that giving this level of control to a single innovator can have deleterious effects on follow-on research.⁴²⁴ While the litigants in the case might not have the incentive to bring these sorts of economic arguments before the court (the patentees would obviously not have any such incentive, and the alleged infringer might own platform patents to which it would hope broad scope would be given in the future), agencies like the DOJ or the FTC could file amicus briefs calling attention to such arguments. If the appellate court enunciated a general policy against allowing platform inventions broad patents, the inferior decisionmakers would of course have to do the fact finding necessary to determine whether the invention at issue was in fact considered a platform invention within the relevant art. With their fact-finding expertise enhanced, however, these decisionmakers would presumably be able to apply the policy with a reasonable degree of success.

The availability of resources like amicus briefs and economic literature notwithstanding, there can be little question that, in absolute terms, courts are significantly handicapped. Nonetheless, because the inquiry must be comparative rather than absolute, courts may still be better than the institutional alternatives. In the case of patents, the two obvious institutional alternatives for policy development are Congress and the PTO. In what follows, I compare the court system to each of these alternatives.

1. *The Comparison to Legislative Policymaking.* — It is beyond dispute that Congress can avail itself of resources that are simply not available to the court system. Innovation does not, however, necessarily lend itself to being regulated by detailed statutes. Because innovation is, almost by definition, unpredictable, predictions embodied in detailed statutes will often be wrong. Moreover, because it is difficult for Congress to revisit its mistakes, these ill-fitting statutes will tend to stay in place and continue to have distorting effects.

423. Sharon Begley & Laura Johannes, *Biotech Firm Says Lilly Drugs Step on Patent*, Wall St. J., June 26, 2002, at B1.

424. See, e.g., Merges & Nelson, *supra* note 164, at 916 (arguing that unduly broad patent scope, in contrast to patent claims more narrowly tailored to inventors' results, makes follow-on research less likely); Rai, *Cumulative Innovation*, *supra* note 160, at 853 (contending that, in context of biopharmaceutical industry, broad patents on upstream research pose primary threat to cumulative innovation).

Consider, for example, one of the few situations in which Congress has added to the patent statute detailed language. In 1995, Congress countered Federal Circuit nonobviousness decisions that appeared to establish a rule against the patenting of chemical processes, even where the materials used in, and produced by, the process were nonobvious, by setting up its own rule. Congress's intricate provision, codified as § 103(b), essentially states that biotechnological processes that use or result in a novel and nonobvious product are always nonobvious.⁴²⁵ This legislation, known as the Biotechnological Process Patents Act of 1995 (BPPA), is problematic on many levels. As an initial matter, nonobviousness is a fact-specific inquiry that cannot be determined through the application of bright-line rules.⁴²⁶ In addition, as discussed further below, the legislation demonstrates how congressional attempts to address questions of patent law and policy often devolve into rent-seeking opportunities for particular industries. For present purposes, however, the more relevant point is that the legislation is simply too detailed to be useful. The statute defines a biotechnological process in terms of the two technologies that have historically been central to biotechnology—recombinant DNA and monoclonal antibody technology. As biotechnology increasingly becomes an information-driven industry, however, this definition of biotechnological process is rapidly becoming outdated.

Congress's attempts at detailed regulation of technology in areas outside the patent law have fared no better. The Semiconductor Chip Protection Act, which sets up a *sui generis* regime of intellectual property protection for chips, was superseded by new technology soon after its creation.⁴²⁷ Commentators have also argued that much of the legislation Congress has adopted to address the new technologies of cyberspace has been shortsighted and ill-fitting.⁴²⁸ Yet this legislation has not been repealed—although in some cases it has been struck down by the Supreme Court as unconstitutional.⁴²⁹

425. 35 U.S.C. § 103(b) (2000).

426. Indeed, in this case, the Federal Circuit was the entity that recognized the difficulties associated with bright-line rules. *In re Ochiai*, a biotechnological process decision issued shortly after the passage of the BPPA, emphasizes that "section 103 requires a fact-intensive comparison of the claimed process with the prior art rather than the mechanical application of one or another *per se* rule." 71 F.3d 1565, 1571 (Fed. Cir. 1995).

427. 17 U.S.C. § 906 (2000); see also A. Samuel Oddi, *An Uneasier Case for Copyright Than for Patent Protection of Computer Programs*, 72 Neb. L. Rev. 351, 450–51 (1993) ("A great deal of time, effort and money was expended . . . to solve what proved to be virtually a nonexistent problem [because] [t]echnological advancements appear to have left whatever copyists there may have been in the lurch.").

428. See, e.g., Suzanna Sherry, *Haste Makes Waste: Congress and the Common Law in Cyberspace*, 55 Vand. L. Rev. 309, 310–17 (2002) (arguing that the Electronic Signatures in Global and National Commerce Act and the Anticybersquatting Consumer Protection Act are shortsighted).

429. The Communications Decency Act, 47 U.S.C. § 223 (2000), which addressed the issue of children's access to Internet pornography, was struck down as unconstitutional in *Reno v. ACLU*, 521 U.S. 844, 849 (1997).

Similarly, suppose Congress responded to the considerable empirical evidence that the economic trajectory of innovation, and hence the role of patents, varies by industry type⁴³⁰ by specifying in detail how antitrust principles, as well as such patentability requirements as nonobviousness and adequate disclosure, should apply in different industries. Suppose further that in the case of the biopharmaceutical industry, where research is currently quite expensive and patents are central to innovation, Congress decided that antitrust principles should have limited scope and that the nonobviousness and disclosure standards should be relatively low.⁴³¹ As matters currently stand, biopharmaceutical innovation is a resource-intensive area. In the future, however, as information technology begins to play a much larger role, biopharmaceutical innovation may become significantly less resource intensive.⁴³² Given the high administrative costs of creating a patent regime for biopharmaceuticals, and the possibility that it may become a poor fit for the industry rather quickly, we are probably better off not having such a regime. By contrast, a common law approach to issues of patent policy, which took account of the disparate economic characteristics of innovation in different industries on a case-by-case basis,⁴³³ could also take account of changes in these characteristics. Although the common law would not necessarily be particularly quick in responding to change, it would probably work more quickly than Congress, which would have to repeal old legislation and enact new legislation on a regular basis.

430. See, e.g., *supra* note 167 (discussing studies focusing on pharmaceutical industry).

431. In offering this hypothetical, I do not mean to suggest that these decisions would necessarily be a good idea as a policy matter. Indeed, a low nonobviousness standard, coupled with a low disclosure standard that allowed for patents of broad scope on upstream innovation, might cause considerable problems for follow-on innovators. See *supra* notes 166, 424 and accompanying text; also cf. Robert M. Hunt, Research Department, Fed. Reserve Bank of Phila., Patentability, Industry Structure, and Innovation 1–2 (Working Paper No. 01-13, Aug. 2001), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=295664 (on file with the *Columbia Law Review*) (setting forth a model of sequential innovation in which rates of innovation are maximized by a high nonobviousness standard for industries that innovate rapidly and a less stringent standard for industries that do not innovate rapidly).

432. See generally Rai, Information Revolution, *supra* note 154 (arguing that potential for genomics to reduce pharmaceutical R&D costs may warrant modification of current pharmaceutical patent regime).

433. See generally Michael A. Carrier, Unraveling the Patent-Antitrust Paradox, 150 U. Pa. L. Rev. 761 (2002) (suggesting a common law approach to the intersection of patent and antitrust that is sensitive to the economic characteristics of innovation in the industry at issue). Factual findings about the economic characteristics of innovation in a particular industry are best considered legislative facts. As contrasted with the adjudicative facts considered in Part I, *supra*, these findings may have relevance across more than one case. Hence, such legislative facts may be appropriate candidates for *de novo* appellate review. By the same token, legislative facts can change. The common law allows legal evolution that reflects factual change.

The alternative of legislation becomes particularly unattractive when one incorporates political economy calculations into the institutional calculus. The standard interest group theory objection to detailed legislative action—that such action is likely to be inefficient because it is subject to undue influence by rent-seeking groups formed around a narrow agenda⁴³⁴—appears to operate with some force in the arena of intellectual property. Even though Congress has generally avoided fleshing out the open-ended language of the patent statute, those amendments that have been made appear to reflect wealth transfers to particular industries. Consider, for example, the aforementioned provisions that set up a per se rule of nonobviousness for certain biotechnological processes.⁴³⁵ Even if nonobviousness could be governed by per se rules (which it cannot), it is not clear why the biotechnology industry should be the sole beneficiary of these rules.

The perils of legislative action in the area of intellectual property can be seen most clearly by looking at copyright law. In copyright law (as contrasted with patent law), Congress has been very active and has created an intricate and dense web of statutory language. The influence of narrowly focused interest groups—primarily content providers of various sorts—has been highly visible. Moreover, with respect to at least some of this legislation, it is difficult to argue that a fair-minded policymaker who had listened to all arguments on how best to promote innovation would have reached the conclusions reached by Congress. Not only has there been a uniform trend towards longer copyright terms,⁴³⁶ but in the context of the recently-enacted Sonny Bono Copyright Term Extension Act (CTEA), content providers managed to persuade Congress to pass legislation extending copyright terms for works that had already been created—and, indeed, many of which would soon fall into the public domain.⁴³⁷ This retroactive copyright term extension legislation is sufficiently anomalous that the Supreme Court agreed to hear *Eldred v. Ashcroft*, which challenged congressional power under the Constitution's copyright and patent clause to pass the extension.⁴³⁸ To be sure, seven justices ultimately rejected the argument that the CTEA did not pass constitutional muster: The majority emphasized that since Congress had, as a historical matter, routinely passed retroactive copyright term extensions, it was dif-

434. See *supra* notes 141–143 and accompanying text.

435. See *supra* notes 425–426 and accompanying text.

436. Brief for Petitioners at 2, *Eldred v. Ashcroft*, 123 S. Ct. 769 (2003) (No. 01-618) (noting that Congress has extended the copyright term eleven times in the last forty years). Although Congress has repeatedly extended the copyright term, two major extensions are particularly salient. In 1976, Congress extended the copyright term for nineteen years. Act of Oct. 19, 1976, Pub. L. No. 94-553, § 304(a), 90 Stat. 2541, 2573–74 (1976). In 1998, it then extended the term by another twenty years. Sonny Bono Copyright Term Extension Act, Pub. L. No. 105-298, § 102, 112 Stat. 2827, 2827 (1998).

437. § 102, 112 Stat. at 2827.

438. *Eldred v. Ashcroft*, 534 U.S. 1126 (2002) (granting writ of certiorari).

ficult to argue that this particular extension was unconstitutional.⁴³⁹ Whatever the merits of the constitutional argument, the overwhelming majority of commentators have noted that extending a copyright term for works that have already been created is very difficult to justify on economic grounds. Indeed, the plaintiffs in the *Eldred* case were supported by an amicus brief filed by seventeen economists, including five Nobel Laureates, who span the ideological spectrum and who represents a virtual “who’s who” of the profession.⁴⁴⁰ Similarly, the Digital Millenium Copyright Act (DMCA) has been roundly criticized for its many specific concessions to rent-seeking copyright owners.⁴⁴¹

2. *The Comparison to PTO Policymaking.* — In the context of developing patent policy, judicial evolution is likely to be superior to legislative action. Whether judicial evolution is superior to administrative action poses a much more difficult question. As noted, courts do not have the resources to conduct complex economic inquiries. By contrast, the modern administrative state was set up precisely for the purpose of regulating an economy that had grown too complex and protean for courts or even Congress to regulate directly. In fact, the PTO has for many years issued guidelines offering evolving interpretations of various elements of the patent statute. In the last several years, for example, it has promulgated for public comment, and subsequently issued, important guidelines on the patentability requirements of utility and written description.⁴⁴² Moreover, the guidelines on utility incorporate, at least implicitly, economic concerns that setting the utility standard too low could impede scientific

439. *Eldred*, 123 S. Ct. 769, 778–80 (2003).

440. Brief of Amici Curiae George A. Akerlof et al. in Support of Petitioners, *Eldred v. Ashcroft*, 123 S. Ct. 769 (2003) (No. 01-618) (signed by, inter alia, Kenneth Arrow, Ronald Coase, and Milton Friedman). As the economists’ brief points out:

[I]t is especially difficult to understand the CTEA’s extension for *existing* works by reference to efficiency. For existing works, particularly those whose pre-CTEA copyrights were about to expire, the social cost of monopoly pricing is at a maximum, and here the extension provides no counter-balancing increase in the incentive to produce new works.

Id. at 11 (emphasis added). The economists’ brief also notes that extending the copyright on existing works is highly unlikely to provide incentives to make investments in, or improvements to, existing work. *Id.* at 9. To the contrary, for everyone other than the copyright owner herself, an extension of the copyright term is likely merely to increase the transaction costs of improvement. *Id.* at 12–15.

441. See, e.g., Yochai Benkler, *Free as the Air to Common Use: First Amendment Constraints on the Enclosure of the Public Domain*, 74 N.Y.U. L. Rev. 354, 421 (1999) (criticizing Congress’s decision in enacting the DMCA to give copyright owners the legal power to “extinguish the user’s privileged uses”); Pamela Samuelson, *Intellectual Property and the Digital Economy: Why the Anti-Circumvention Regulations Need to Be Revised*, 14 Berkeley Tech. L.J. 519, 523 (1999) (“[B]y colorful use of its high rhetoric and forceful lobbying, Hollywood and its allies were successful in persuading Congress to adopt the broad anti-circumvention legislation they favored . . .”). With the DMCA, however, it may be more difficult to argue unequivocally that no fair-minded policymaker could believe it promotes the goals of intellectual property law.

442. See *supra* note 160.

progress by creating a transaction-cost-heavy thicket of patents on basic research.⁴⁴³ These concerns were voiced by a variety of scientific research groups, including perhaps most prominently the scientists at the National Institutes of Health (NIH).⁴⁴⁴

Congress has not, however, given the PTO substantive rulemaking authority. For this reason, under the Supreme Court's recent applications of the court/agency deference doctrine first enunciated in the 1984 case *Chevron U.S.A. Inc. v. Natural Resources Defense Council*,⁴⁴⁵ interpretations of the patent statute issued in the PTO guidelines are not entitled to deference. According to the Court's recent decisions, interpretations contained in agency guidelines and regulations merit *Chevron* deference only if they are issued by an agency with substantive rulemaking authority.⁴⁴⁶

Moreover, there are reasons to be wary about granting the PTO substantive rulemaking authority. As matters currently stand, the PTO does not have economic expertise within the agency. In particular, as con-

443. Cf. Heller & Eisenberg, *supra* note 160, at 698 ("The result [of a regime of expansive patent rights] has been a spiral of overlapping patent claims in the hands of different owners, reaching ever further upstream in the course of biomedical research.").

444. Admittedly, the PTO did not go as far as NIH would have liked. See David Dickson, NIH Opposes Plans for Patenting 'Similar' Gene Sequences, *Nature*, May 2000, at 3 (discussing NIH pressure on PTO to raise utility standard but noting its opposition to PTO decision to allow assertions of utility based on homology to known sequences).

445. 467 U.S. 837 (1984). *Chevron's* famous two-step procedure requires the reviewing court first to determine whether the statutory language that is being interpreted by the agency is unambiguous with respect to the question at issue. If it is, "that is the end of the matter; for the court, as well as the agency, must give effect to the unambiguously expressed intent of Congress." *Id.* at 842-43. If the court determines that the statute is ambiguous with respect to the question at issue, it proceeds to step two, which requires that the court defer to the agency's construction, so long as that construction is reasonable. *Id.* at 843. In the *Chevron* decision itself, see *id.* at 843-44, as well as in subsequent decisions applying *Chevron*, the Court has justified this deferential framework by arguing that where Congress leaves an ambiguity in a statute, it is effectively delegating to the agency that administers the statute the authority to resolve that ambiguity. See, e.g., *FDA v. Brown & Williamson Tobacco Corp.*, 529 U.S. 120, 132 (2000); *United States v. Haggar Apparel Co.*, 526 U.S. 380, 392-93 (1999); *Adams Fruit Co. v. Barrett*, 494 U.S. 638, 649 (1990).

446. See *EEOC v. Arabian Am. Oil Co.*, 499 U.S. 244, 257 (1991) (holding that EEOC guidelines interpreting Title VII statutory language do not deserve deference because EEOC does not have authority to promulgate rules or regulations under Title VII). Moreover, according to the Court's 2001 decision in *United States v. Mead Corp.*, even rules or regulations issued by an agency with substantive rulemaking authority will only merit deference if they are "promulgated in the exercise of that authority" (e.g. through notice-and-comment rulemaking). 533 U.S. 218, 227 (2001). For an exhaustive discussion on the Court's gradual narrowing of the scope of *Chevron*, see Thomas W. Merrill & Kristin E. Hickman, *Chevron's Domain*, 89 Geo. L.J. 833, 840-48, 856-58 (2001). This narrowing trend has been championed by commentators who argue against a general delegation of interpretative authority to agencies on the grounds that the language of the Administrative Procedure Act requires courts to decide legal questions. See, e.g., John F. Duffy, *Administrative Common Law in Judicial Review*, 77 Tex. L. Rev. 113, 189-210 (1998) (arguing that language, structure, and legislative history of APA indicate congressional intent that courts interpret laws without automatic deference to agencies).

trusted with most other agencies, the PTO does not employ *any* economists.⁴⁴⁷ Thus it is by no means clear that the PTO has any greater institutional resources on matters of innovation policy than do courts. In addition, notwithstanding its recent response to research groups opposed to the patenting of basic science, the PTO appears to have developed an institutional culture that treats patentees as “clients” to be served rather than as claimants who must present a case for being entitled to a patent.⁴⁴⁸ Even with the implementation of an opposition proceeding and the elimination of monetary incentives to grant patents, the extent to which this culture will change is not clear. At a minimum, then, any grant of substantive rulemaking authority should be premised on very substantial reform of the agency: This should include not only reform of the process by which individual patents are granted but also wholesale reform of the policymaking apparatus within the agency.⁴⁴⁹

The likelihood that Congress will enact wholesale reform of the PTO’s policymaking apparatus appears slim, however. Patent holders are likely to oppose vigorously legislation that endows the PTO with economic expertise. Such legislation would presumably lead to the presence within the PTO of at least some economists who are skeptical of the position that strong patents are always the best mechanism for promoting innovation.⁴⁵⁰

447. Interview with Stephen Kunin, PTO Deputy Commissioner for Patent Policy, in Seattle, Wash. (July 20, 2002).

448. As Cecil Quillen, a former general counsel at Eastman Kodak Company, recently noted:

There has long been a symbiotic relationship between the Patent Office and those who practice before it, and those who litigate its results. Each has depended on the other for their livelihoods. Historically the way it worked was that the Patent Office would issue a few more patents each year, which would require a few more patent applications, which would require a few more patent attorneys and patent examiners, and on and on and on.

Quillen, *supra* note 409, at 17.

449. If the opposition-proceeding reform proposed in this Article were implemented, there might be one context in which PTO interpretation would be entitled to deference. According to the Court’s decision in *United States v. Mead Corp.*, agency interpretations rendered during adjudication are entitled to deference if the adjudication involves a public, trial-type procedure. 533 U.S. at 229–30. Thus, to the limited extent the PTO could make general legal or policy interpretations in the context of a fact-centered opposition proceeding, these interpretations might be given *Chevron* deference. For the reasons discussed in the text, however, even deference in this narrow context may not be wise as a policy matter.

450. See *supra* notes 353–354 and accompanying text (discussing traditional economic view that recognizes both the costs and benefits of patents in promoting innovation). By way of contrast, it should be possible to persuade Congress to implement PTO reform—for example, opposition proceedings—that improves the agency’s fact finding abilities. Because patent holders have some interest in being able to defeat specific patents held by their adversaries, they may not mount vigorous opposition to the implementation of opposition proceedings. See *supra* text accompanying notes 186–187.

CONCLUSION

Given the interdependence of the multiple institutional problems in the U.S. patent system, reform of this system must be multi-institutional. Moreover, such reform must focus on remedying the system's substantial deficiencies in the areas of fact finding, policy application, and policymaking. Under a traditional administrative law model, primary responsibility for questions of fact would rest with the PTO. The role of the courts—whether the trial court or the appellate court—would be secondary. With respect to fact finding, adopting certain aspects of the traditional administrative model is clearly necessary. Thus Congress should bring the PTO into the twenty-first-century administrative state by endowing the agency with the expertise and neutrality necessary for it properly to wield significant fact-finding responsibility. Most importantly, opposition proceedings that give the PTO greater access to relevant industry knowledge should be implemented. With respect to commercially significant patents, such proceedings would address the information deficit necessarily faced by the PTO when it evaluates patent applications. Patents that had undergone opposition proceedings might actually merit the high presumption of validity that the Federal Circuit currently accords to all patents. In addition, the fee structure of the PTO should be altered so as to give PTO examiners no greater incentives to deny patents than to issue them.

At the same time, however, giving plenary responsibility on factual questions to the PTO would not be cost effective. Accordingly, the trial courts should also be endowed with fact-finding expertise. Specifically, one or more district courts should be designated to take all patent cases. These district courts should also be given the resources for hiring court-appointed expert witnesses, special masters, and technical advisors who have skill in the particular science or technology relevant to the patent case. In addition, consistent with the Seventh Amendment, the role of the trial court jury can and should continue to be restricted. With greater fact-finding expertise, the PTO and particularly the trial courts would presumably also be better able to apply policy-driven standards that require fact finding.

Solving the allocation puzzle for questions of policymaking is more challenging. Each of the available institutional options—the legislature, the PTO, and the courts—has substantial associated liabilities. The most disadvantaged institutional actor is probably Congress. The cumbersome machinery of detailed legislation is ill-suited for the dynamic nature of innovation. Moreover, in the area of intellectual property, Congress appears particularly susceptible to influence by rent-seeking entities. The institutional choice thus comes down to one between the PTO and the appellate courts. On first examination, given the Federal Circuit's apparent aversion to policy analysis, its possible tendency towards pro-patent bias, and the well-known difficulties courts face in undertaking policymaking, the PTO might appear the better option. In order for the

PTO to be a reasonably good policymaker, however, Congress would have to change the PTO into an altogether different agency: Merely implementing reform that improved the agency's fact-finding capabilities would not be sufficient. By contrast, the level of change required in the court system would be less substantial. What we would need is greater generalist input in the appellate process. This generalist input should come from several sources. Amicus briefs from agencies like the Federal Trade Commission that have strong traditions of economic analysis could assist courts with their policymaking. The Supreme Court should increase its review of cases that raise questions of substantive patent policy. Ideally, Congress would also implement a system of generalist en banc review for the Federal Circuit: Under such a system, judges from other circuits would, from time to time, review Federal Circuit decisions.

Despite the well-theorized institutional shortcomings of courts when it comes to policymaking, in the specific area of patent law, policy development through the court system is probably the best of the available options. The bottom line, then, is that the substantive failures of the patent system need not persist. By paying attention to institutional design and revising our institutions accordingly, we can achieve the patent system we should have had all along.