
From the SelectedWorks of Herbert Hovenkamp

August 2009

IP and Antitrust: Errands into the Wilderness

Contact
Author

Start Your Own
SelectedWorks

Notify Me
of New Work



Available at: http://works.bepress.com/herbert_hovenkamp/8

IP AND ANTITRUST: ERRANDS INTO THE WILDERNESS

Christina Bohannon*
Herbert Hovenkamp**

Introduction

On May 11, 1670, Puritan Congregational Minister Samuel Danforth delivered a sermon to the Massachusetts Colonial General Assembly. Its purpose was to solicit God's blessing on the upcoming public election, and Danforth's topic was New England's Errand into the Wilderness.¹ This forced Errand, begun half a century earlier, was intended both to punish and to rehabilitate. The English Church had become corrupt and its theology misguided. God then declared that the true church needed some time out and selected a few hundred people for the extreme harshness of New England winters where they could purify themselves and rebuild. Danforth lamented the fact that only fifty years after the errand began, signs of rationalism and sophistry were already re-appearing in New England religious thought. Although Danforth was pessimistic about these developments, others came to celebrate American exceptionalism in several areas, including religion, politics, and economics.

Our topic here is some other Errands into the Wilderness—one largely completed and others unfinished—that are essential to innovation, competition, and economic growth. The errand that is largely completed is antitrust's decades-long period of isolation, redefinition and retrenchment. The errands that are not yet finished belong to the IP laws, particularly patent and copyright, the two bodies of IP law expressly authorized by the Intellectual Property Clause of the Constitution.² By the 1970s antitrust had lost much of its concern with economic competition and begun protecting less efficient small businesses from the lower costs of larger firms. Then, beginning in the 1970s the Supreme Court changed course, refocusing antitrust law on the protection of consumers. By the same token, patent and copyright law have lost their focus on facilitating the type and amount of innovation needed to benefit consumers, and turned toward the protection of rights holders.

The linkage of intellectual property and competition policy is hardly novel.³ But our

* Associate Professor of Law, University of Iowa College of Law. The authors thank the Ewing Marion Kauffman Foundation for its generous financial support. We also appreciate helpful comments from Professors John Golden and Mark Lemley, from participants in the Colloquium on innovation policy at NYU law school lead by Professors Rochelle Dreyfuss and Harry First, where most of these ideas were initially presented, March, 2009; and also the participants in a faculty seminar at Notre Dame University Law School, April 2009.

** Ben V. & Dorothy Willie Professor, University of Iowa College of Law.

¹ Samuel Danforth, Pastor, Church of Christ in Roxbury N.E., *A Brief Recognition of New Englands Errand into the Wilderness* (May 11, 1670), <http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1038&context=libraryscience>. The phrase "Errand into the Wilderness" also became the title of a famous book by the distinguished Colonial historian Perry Miller. PERRY MILLER, *ERRAND INTO THE WILDERNESS* (Harvard Univ. Press 1956) (1952).

² See U.S. CONST. art. I, § 8, cl. 8.

³ See, e.g., HERBERT HOVENKAMP, MARK D. JANIS, MARK A. LEMLEY AND CHRISTOPHER LESLIE, *IP AND*

concern here reaches beyond applying established antitrust doctrine in IP-intensive areas. Rather, we try to develop the basis for a common legal theory for fostering innovation and growth. Identifying the appropriate scope of intellectual property protection is as much a question of competition policy as of patent or copyright policy. Overly broad IP rights or infringement doctrine that is too lax about proof of harm serves to create unjustified regions that are protected from competition and that threaten innovation. In very much the same way, excessive enforcement of the antitrust laws without concern about competitive injury protects firms from competition and undermines the incentive to innovate. In both cases consumers are the victims.

In both antitrust and IP, meaningful reform requires a new focus on the degree and kind of harm required for a violation. As the Supreme Court has discovered through many years of interpreting the antitrust laws in private plaintiff actions, one way to avoid the pitfalls created by ambiguous and complex substantive laws is to ensure that the harm that is claimed is consistent with the underlying purpose of those laws.⁴ For example, predicting the economic consequences of a merger is extraordinarily difficult and the chance of an error correspondingly high.⁵ But if a plaintiff is complaining that a merger caused *more* rather than less competition in a market, why bother with the difficult substantive analysis? The very nature of the plaintiff's claim tells us that we do not want to condemn this particular merger, at least not for the reason that the plaintiff claims. As we argue in Part V, we could go a long way toward correcting problems of IP overreaching by avoiding the truly difficult decisions that are needed to formulate substantive patent and copyright rules. Rather, courts need to ask more frequently whether the type of harm of which the plaintiff complains is sufficiently related to the underlying goals of those laws, which can be defended only on the grounds that they encourage innovation.

We begin in Part I by giving an account of antitrust's errand into the wilderness, and of how reform was accomplished largely by the judiciary and in apparent conflict with a statute that seemed both clear and inflexible. Then Part II examines the state of reform in the IP laws generally, and turns to specific issues involving patents in Part III and copyright in Part IV. Finally, in part V we urge courts to develop a concept of "IP injury," similar to the concept of "antitrust injury" in the antitrust laws, which links the type of harm that a plaintiff must show to the underlying purpose of those laws.

Our evaluation of these three bodies of law that are so central to innovation and competition also leads us to conclude that courts are far more likely than Congress to be the engines of true

ANTITRUST: AN ANALYSIS OF ANTITRUST PRINCIPLES APPLIED TO INTELLECTUAL PROPERTY LAW (2d ed. 2010) (in press).

⁴ See discussion *infra*, text at notes 274-279.

⁵ On the need for simplifying assumptions and prophylactic rules in merger analysis, *see* 4 PHILLIP E. AREEDA & HERBERT HOVENKAMP, ANTITRUST LAW ¶905 (3d ed. 2009).

reform. The revolution in antitrust was undertaken almost exclusively by the courts, in the face of a clear and aggressive statute that would seem to make a strenuous harm requirement impossible. The transformation that is currently occurring in patent law is also very largely the work of the Supreme Court. While the cleansing of the Copyright Act remains mainly in the future, nothing gives us hope that the reforms will come from Congress.

I. Antitrust's Errand into the Wilderness

The story of antitrust's Errand into the Wilderness is well known. The half-century period that ended in the late 1970s had seen many antitrust infidelities, mainly from expansion that today seems unprincipled, given that injury to competition was so often absent. Among these infidelities was the doctrine that mergers in highly competitive markets should be condemned if they permitted the post-merger firm to undersell smaller rivals;⁶ the *Utah Pie* doctrine that recent entrants into a market could not use aggressive pricing against a dominant firm there if they were charging higher prices somewhere else;⁷ that vertical nonprice restraints should be unlawful per se;⁸ that tying arrangements should be unlawful even if the seller lacked serious market power and the tie excluded no one;⁹ and that purely vertical agreements maintaining either maximum or minimum resale prices were unlawful per se.¹⁰ In this period, antitrust law was substantially overdeterrent. That is, it condemned a good deal of conduct that was procompetitive and in the process forced consumers to pay higher prices.

What went wrong with antitrust following World War II was not all that different from what is wrong with the intellectual property laws today. In both areas of law, the legal policy became

⁶ *Brown Shoe Co. v. United States*, 370 U.S. 294 (1962) (condemning a merger because the post-merger firm would be able to undersell smaller stores); *United States v. Von's Grocery Co.*, 384 U.S. 270 (1966) (similar); *FTC v. Procter & Gamble Co.*, 386 U.S. 568 (1967) (condemning a merger because it would create economies in advertising).

⁷ *Utah Pie Co. v. Cont'l Baking Co.*, 386 U.S. 685 (1967) (unlawful for firms operating in many markets to undersell a dominant firm in its local market).

⁸ *United States v. Arnold, Schwinn & Co.*, 388 U.S. 365 (1967) (vertical nonprice restraints unlawful per se), *overruled by Cont'l T. V., Inc. v. GTE Sylvania Inc.*, 433 U.S. 36 (1977) (vertical nonprice restraints to be subjected to rule of reason).

⁹ *Int'l Salt Co. v. United States*, 332 U.S. 392 (1947) (tying of patented product presumptively unlawful), *overruled by Ill. Tool Works Inc. v. Indep. Ink, Inc.*, 547 U.S. 28 (2006) (no presumption of market power from patented tying product).

¹⁰ *Dr. Miles Med. Co. v. John D. Park & Sons Co.*, 220 U.S. 373 (1911) (minimum resale price maintenance unlawful per se), *overruled by Leegin Creative Leather Prods., Inc. v. PSKS, Inc.*, 127 S. Ct. 2705 (2007) (minimum resale price maintenance addressed under rule of reason); *Albrecht v. Herald Co.*, 390 U.S. 145 (1968) (maximum resale price maintenance unlawful per se), *overruled by State Oil Co. v. Khan*, 522 U.S. 3 (1997) (maximum resale price maintenance to be addressed under rule of reason).

disconnected from its articulated goals and began pursuing other ends. Antitrust shielded small businesses from competition that would have benefitted consumers. Similarly, IP law expanded entitlements for the benefit of patent and copyright holders, often at the expense of innovation, which always relies on the work of predecessors as well as a robust public domain. The ultimate victims, once again, were consumers.

In the case of antitrust, people initially placed the blame for this disconnection on the Supreme Court—something that Robert Bork and Ward Bowman did so forcefully in their disturbing 1960s article on “The Crisis in Antitrust,”¹¹ and that Bork expanded a decade later in *The Antitrust Paradox*.¹² More recently, the tendency has been to blame private plaintiffs and the private treble damages enforcement system.¹³ In reality, however, Congress and the government antitrust enforcement agencies were behind most of it. For example, *Schwinn*¹⁴ was an aggressive decision that condemned manufacturer-imposed territorial restrictions because they were restraints on alienation, not because they had any impact on competition. The *Brown Shoe* decision condemned a merger because it injured smaller rivals, albeit by reducing prices to consumers.¹⁵ But the plaintiffs in these cases were not *Pop’s Bike and Trike* or *Sally’s Little Shoe Shop*. In both cases, as well as other big Supreme Court decisions of this vintage, the plaintiff was the United States government. The Supreme Court rarely did more than give the government what it asked for. Yes, the Supreme Court did condemn many mergers precisely because they created efficiencies that might injure competitors, but it did so, at least in large part, because that is what the Antitrust Division and the Federal Trade Commission wanted.¹⁶

For example, speaking through Archibald Cox as Solicitor General, the government identified low consumer prices as the primary evil brought about by the Brown Shoe/Kinney merger:

¹¹ Robert H. Bork & Ward Bowman, Jr., *The Crisis in Antitrust*, 65 COLUM. L. REV. 363 (1965).

¹² ROBERT H. BORK, *THE ANTITRUST PARADOX: A POLICY AT WAR WITH ITSELF* (1978).

¹³ See, e.g., William J. Baumol & Janusz A. Ordover, *Use of Antitrust to Subvert Competition*, 28 J.L. & ECON. 247 (1985) (antitrust protects inefficiency for benefit of small business); Frank H. Easterbrook, *Predatory Strategies and Counterstrategies*, 48 U. CHI. L. REV. 263 (1981) (similar); Edward A. Snyder & Thomas E. Kauper, *Misuse of the Antitrust Laws: The Competitor Plaintiff*, 90 MICH. L. REV. 551 (1991) (similar).

¹⁴ *United States v. Arnold, Schwinn & Co.*, 388 U.S. 365 (1967), *overruled by* *Cont’l T. V., Inc. v. GTE Sylvania Inc.*, 433 U.S. 36 (1977).

¹⁵ *Brown Shoe Co. v. United States*, 370 U.S. 294 (1962) (government suit to break up a merger).

¹⁶ In addition to *Brown Shoe*, see *FTC v. Procter & Gamble Co.*, 386 U.S. 568 (1967) (condemning a product-extension merger, a type of conglomerate acquisition, because it would lead to economies in marketing that would enable P&G to undersell rivals). Other decisions are discussed in 4 PHILLIP E. AREEDA & HERBERT HOVENKAMP, *ANTITRUST LAW* ch. 9 (3d ed. 2009), and 4 *id.*

[T]he integration of manufacturer-retailer Brown with the large Kinney retail organization will seriously aggravate the difficulties that independent retailers are already having in competing with the substantial and ever-expanding retail chains. The manufacturer-owned or controlled retail outlet can sell its own product at a significantly lower price than the nonintegrated independent retailer can obtain for a comparable product The conclusion was inevitable that the advantages the merged company would have over its smaller retailing competitors would be so great as to threaten to become decisive.¹⁷

Even as early as the *International Salt* case, the Court granted the government's request to condemn a tying arrangement without any showing of either market power in the tying product or significant harm in the market for the tied-up salt. That is, the government showed no injury to competition whatsoever but believed it was wrong for a patentee to insist that users of its patented machine also use its own salt, a common commodity not capable of being monopolized.¹⁸

As for private plaintiffs, they assuredly have a tendency to push the envelope, but in most subsequent private actions they asked the federal courts to do no more than give them what had already been given to the government. For example, the plaintiffs in *Brunswick Corp. vs. Pueblo Bowl-O-Mat, Inc.*¹⁹ had simply requested that the court follow *Brown Shoe* by protecting their small bowling alley from a rival that threatened to become more robust and efficient as a result of a merger. The Third Circuit obliged with a detailed analysis of the *Brown Shoe* decision. The court found that the main difference in the present case was that the bowling alley market in Pueblo, Colorado, was far more concentrated than the shoe market in *Brown*.²⁰ In reversing, the Supreme Court created the "antitrust injury" doctrine, discussed below, which stood the rationale of *Brown Shoe* on its head.²¹

Dating the beginning of antitrust's Errand into the Wilderness is difficult. It may have been 1957, when Ward Bowman exploded the traditional leverage theory of tying arrangements, which had suggesting that tying of a monopoly product to a competitive product could turn one monopoly

¹⁷ Brief for the United States at 48, *Brown Shoe*, 370 U.S. 294 (No. 4).

¹⁸ *Int'l Salt Co. v. United States*, 332 U.S. 392 (1947) (accepting government's argument that power and anticompetitive effects be presumed from existence of a patent), *overruled by* *Ill. Tool Works Inc. v. Indep. Ink, Inc.*, 547 U.S. 28 (2006) (rejecting patent presumption).

¹⁹ 429 U.S. 477 (1977).

²⁰ *NBO Indus. Treadway Cos. v. Brunswick Corp.*, 523 F.2d 262 (3d Cir. 1975) (permitting plaintiff to challenge efficiency creating merger that injured it).

²¹ *See infra* text accompanying notes 267–72.

into two.²² Perhaps it was 1958 when John McGee debunked the theory that Standard Oil had engaged in predatory pricing,²³ or 1969 when the Neal Report on the antitrust laws was released, provoking a sharp reaction that rejected its interventionist recommendations.²⁴ The Errand may have begun in 1975, when Phillip E. Areeda and Donald F. Turner published their law review article advocating strict, cost-based standards for predatory pricing,²⁵ or 1978 when they published the first three volumes of the *Antitrust Law* treatise.²⁶ Or maybe it began in 1976 when then professor Richard A. Posner published the first edition of his monograph on *Antitrust Law*,²⁷ or 1978 when Robert Bork popularized Chicago School ideas in his influential book, *The Antitrust Paradox*.²⁸ Perhaps a better candidate is 1981, when President Ronald Reagan appointed William F. Baxter as head of the Antitrust Division. In any event, antitrust's errand began largely in academic literature and then moved into government enforcement and the courts.

Today the antitrust landscape differs so much from the view of *Brown Shoe* that it could barely be recognized from that vantage point. Five years after *Brunswick*, the Supreme Court imposed significant restrictions on antitrust standing that limit private enforcement to people who

²² Ward S. Bowman, Jr., *Tying Arrangements and the Leverage Problem*, 67 YALE L.J. 19 (1957) (debunking notion that tying a monopoly product to a competitive product enables a firm to turn one monopoly into two). For a recent attempt to revitalize the leverage theory, see Einer Elhauge, *Tying, Bundled Discounts, and the Death of the Single Monopoly Profit Theory*, 123 HARV. L. REV. (forthcoming Dec. 2009) (manuscript available at <http://ssrn.com/abstract=1345239>). For a reply, see Erik N. Hovenkamp and Herbert Hovenkamp, *Tying, Price Discrimination, and Antitrust Policy* (Iowa Legal Studies Working Paper, Aug. 2009), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1443284.

²³ John McGee, *Predatory Price Cutting: The Standard Oil (N.J.) Case*, 1 J.L. & ECON. 137 (1958) (arguing that Standard Oil did not engage in predatory pricing, but was simply more efficient than its rivals). McGee's work was recently called into question in James A. Dalton and Louis Esposito, *Predatory Price Cutting and Standard Oil: A Re-examination of the Trial Record*, 22 RES. L. & ECON. 155 (2007) (re-examining the trial record and finding numerous instances of predation, in conflict with McGee's conclusions). See also RON CHERNOW, *TITAN: THE LIFE OF JOHN D. ROCKEFELLER, SR.* 113–17, 144–47, 202–25, 251–58 (1998) (documenting various rebate agreements between Standard and the railroad).

²⁴ 115 CONG. REC. 13890 (1969); see Herbert Hovenkamp, *The Neal Report and the Crisis in Antitrust*, COMPETITION POL'Y INT'L, Spring 2009, at 217 (chronicling harsh reaction to Report that advocated aggressive use of antitrust laws).

²⁵ Phillip Areeda & Donald F. Turner, *Predatory Pricing and Related Practices Under Section 2 of the Sherman Act*, 88 HARV. L. REV. 697 (1975) (advocating restrictive cost-based tests for predatory pricing)

²⁶ PHILLIP E. AREEDA & DONALD F. TURNER, *ANTITRUST LAW* (1978) (advocating more restrictive rules for antitrust enforcement, particularly by private plaintiffs)

²⁷ RICHARD A. POSNER, *ANTITRUST LAW: AN ECONOMIC PERSPECTIVE* (1976) (similar).

²⁸ BORK, *supra* note 12.

suffered demonstrable injury as a result of decreased competition.²⁹ In two important decisions twenty years apart, the Supreme Court greatly strengthened pleading and proof requirements, imposing harsh standards for summary judgment in its 1986 *Matsushita* decision³⁰ and strict pleading standards in its 2007 *Twombly* decision.³¹ On substantive issues, the Supreme Court has made exclusionary practices much more difficult to prove, with holdings on predatory pricing in *Brooke Group* in 1993³² and *Weyerhaeuser* in 2007³³ that imposed strict cost-based pricing tests as well as requiring a market capable of profitable monopolization. These rules largely followed the Areeda and Turner recommendations of 1975 and 1978. In 2004, the Supreme Court considerably narrowed the law of unilateral refusals to deal by dominant firms.³⁴ It has completely rewritten the law of vertical restraints, removing such common law concerns as restraints on alienation, which have nothing to do with competition policy.³⁵ More generally, the courts have greatly reduced the use of per se rules, with their automatic inference of competitive harm, except in cases of naked collusion. Rather, they have required plaintiffs to prove market power and anticompetitive effects.³⁶

²⁹ *Assoc. Gen. Contractors of Cal., Inc. v. Cal. State Council of Carpenters*, 459 U.S. 519, 542–44 (1983) (restrictive test for antitrust standing to sue); see 2A PHILLIP E. AREEDA & HERBERT HOVENKAMP, *ANTITRUST LAW* ¶ 335d (3d ed. 2008).

³⁰ *Matsushita Elec. Indus. Co. v. Zenith Radio Corp.*, 475 U.S. 574 (1986) (adopting strict standard for summary judgment in antitrust case); see 2 PHILLIP E. AREEDA & HERBERT HOVENKAMP, *ANTITRUST LAW* ¶ 308 (3d ed. 2007).

³¹ *Bell Atl. Corp. v. Twombly*, 550 U.S. 544 (2007) (adopting strict standard for pleadings alleging unlawful antitrust conspiracy); see 2 AREEDA & HOVENKAMP, *supra* note 30, ¶ 307 (3d ed. 2007).

³² *Brooke Group Ltd. v. Brown & Williamson Tobacco Corp.*, 509 U.S. 209 (1993) (adopting strict rules for predatory pricing cases, following 3 PHILLIP E. AREEDA & HERBERT HOVENKAMP, *ANTITRUST LAW* ¶¶725-728 (3d ed. 2008)).

³³ *Weyerhaeuser Co. v. Ross-Simmons Hardwood Lumber Co., Inc.*, 549 U.S. 312 (2007) (adopting strict cost based rules for predator purchasing, following 3 AREEDA & HOVENKAMP, note 32 at 747).

³⁴ *Verizon Commc'ns Inc. v. Law Offices of Curtis V. Trinko, LLP*, 540 U.S. 398 (2004) (narrowing scope of antitrust law on unilateral refusals to deal, following 3 Areeda & Hovenkamp, note 30 at ¶¶771-774).

³⁵ In chronological order: *Cont'l T. V., Inc. v. GTE Sylvania Inc.*, 433 U.S. 36 (1977) (applying rule of reason to vertical nonprice restraints); *Jefferson Parish Hosp. Dist. No. 2 v. Hyde*, 466 U.S. 2 (1984) (tightening up standards for tying arrangements); *State Oil Co. v. Khan*, 522 U.S. 3 (1997) (applying rule of reason to maximum price fixing); *Ill. Tool Works Inc. v. Indep. Ink, Inc.*, 547 U.S. 28 (2006) (overruling presumption that patented tying product confers sufficient market power to make a tie unlawful); *Leegin Creative Leather Prods., Inc. v. PSKS, Inc.*, 127 S. Ct. 2705 (2007) (applying rule of reason to resale price maintenance).

³⁶ *E.g.*, *Cal. Dental Ass'n. v. FTC*, 526 U.S. 756 (1999) (adopting lenient rules for evaluating price-affecting joint conduct)

II. IP: Errands That Have Barely Begun

Intellectual property law today is in a place closely resembling antitrust policy in the 1960s. As with antitrust, the argument that IP law has become too detached from its central concern with protecting innovation has begun in the academic literature and is fully underway. In patent law, the courts are beginning to respond; in copyright, not so much.

The *articulated* goal of intellectual property law is economic, just as it is for antitrust. As the Supreme Court has recognized, the “economic philosophy” of the Constitution’s Intellectual Property Clause³⁷ is to “advance public welfare” through the patent and copyright systems.³⁸ This is similar to saying that the goal of the antitrust laws is to advance the public welfare by promoting competition. Competition and innovation are two blades of the same scissors. Traditional competition improves economic welfare in the shorter run by keeping markets competitive, which means that prices are reasonably close to cost and there are no unnecessary restraints on output or entry. Innovation policy pursues the same welfare goal, but focuses on more “dynamic” concerns, by giving people proper incentives to develop new ideas and technologies that society will value.³⁹ In both cases the protected class is consumers, who benefit from lower prices and improved products and services.

A. The Articulation of Economic Goals

While antitrust rules are more explicitly grounded in economics than intellectual property rules are, the latter play at least as important a role in economic growth. It has been clear since the work of Joseph Schumpeter, later elaborated in Robert W. Solow’s work on the neoclassical growth model, that the economic gains from innovation dwarf those from capital accumulation and increased price competition.⁴⁰ An important but often overlooked corollary of this proposition is that restraints on innovation can do far more harm to the economy than restraints on traditional output or pricing.⁴¹

³⁷ U.S. CONST. art. I, § 8, cl. 8 (“Congress shall have Power To . . . promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries.”).

³⁸ *Mazer v. Stein*, 347 U.S. 201, 219 (1954) (“The economic philosophy behind the clause empowering Congress to grant patents and copyrights is the conviction that encouragement of individual effort by personal gain is the best way to advance public welfare through the talents of authors and inventors in ‘Science and useful Arts.’”).

³⁹ See Herbert Hovenkamp, *Restraints on Innovation*, 29 CARDOZO L. REV. 247 (2007) (advocating more aggressive use of antitrust laws against innovation reducing restraints).

⁴⁰ See JOSEPH SCHUMPETER, *CAPITALISM, SOCIALISM AND DEMOCRACY*, chs. 7-8 (1942); Robert M. Solow, *A Contribution to the Theory of Economic Growth*, 70 Q.J. ECON. 65 (1956); Robert M. Solow, *Technical Change and the Aggregate Production Function*, 3 REV. ECON. & STAT. 312 (1957); Herbert Hovenkamp, *Schumpeterian Competition and Antitrust*, COMPETITION POL’Y INT’L, Autumn 2008, at 273.

⁴¹ See Hovenkamp, *supra* note 39.

Nevertheless, the concern with economic welfare is much more direct and appears much more frequently in antitrust case law than in IP case law. In antitrust cases, courts often state that promoting economic welfare through competition is the singular goal. One rarely sees an appellate antitrust opinion that reaches the merits and does not say something about whether the practice under examination undermines competition by raising prices above cost, reducing output, injuring consumers, excluding rivals, and the like. In fact, the antitrust injury doctrine requires precisely that in private actions.⁴² Most lower court IP cases are less explicit. Rather than looking to first principles to determine the best way to promote public welfare through innovation, courts in IP cases generally focus on interpreting particular statutory provisions. The Supreme Court and the Federal Circuit in patent cases certainly address fundamental concerns about innovation incentives more frequently than other courts do, although mainly in decisions that interpret the relationship between the IP laws and the Constitution.⁴³

One consequence of the antitrust revolution has been the gradual suppression of noneconomic goals. Over its history, a number of goals have been articulated for the antitrust laws, including a guarantee of “fairness” or justice in business practices, the protection of small business, or some form of populism.⁴⁴ Today, however, all sides of the antitrust debate seem to agree that the antitrust laws are designed to further some version of economic competition.⁴⁵ To be sure, important differences remain. Some prefer an articulation of economic competition that maximizes total economic welfare, which is the sum of producer and consumer wealth. This is the view most consistent with neoclassical economics generally, whose concern is to maximize the size of the pie without regard to how resources are distributed. A softer variation is “consumer welfare,” which seeks to maximize the size of consumers’ surplus. Those holding this view are willing to condemn

⁴² See *infra* text accompanying note 274-278.

⁴³ *E.g.*, *KSR Int’l Co. v. Teleflex Inc.*, 550 U.S. 398 (2007) (nonobviousness; discussing relevance of Constitution’s IP Clause); *Eldred v. Ashcroft*, 537 U.S. 186 (2003) (constitutionality of statutory amendment extending copyright term retroactively; First Amendment and IP Clause); *Graham v. John Deere Co.*, 383 U.S. 1 (1966) (nonobviousness; IP clause); *Sony Corp. v. Universal City Studios, Inc.*, 464 U.S. 417 (1984) (First Amendment and IP Clause).

⁴⁴ See, *e.g.*, Robert H. Lande, *Wealth Transfers as the Original and Primary Concern of Antitrust: The Efficiency Interpretation Challenged*, 34 HASTINGS L.J. 65 (1982) (consumer welfare most prominent concern of antitrust laws); Louis B. Schwartz, “Justice” and Other Non-Economic Goals of Antitrust, 127 U. PA. L. REV. 1076 (1979) (antitrust concerned with leveling playing field between large and small business). For evaluations, see HERBERT HOVENKAMP, FEDERAL ANTITRUST POLICY: THE LAW OF COMPETITION AND ITS PRACTICE ch. 2 (3d ed. 2005); BORK, *supra* note 12, at 5–8.

⁴⁵ See HOVENKAMP, *supra* note 44, at ch. 2.

a practice that harms consumers, even though it benefits producers by a larger amount.⁴⁶ The classic example is the merger that results in higher consumer prices, but produces an even larger efficiency gain to the merging parties. Assuming no one else is affected, such a merger is “efficient” in the neoclassical sense because total value is increased, and we don’t care about who is richer. However, the consumer welfare argument goes, antitrust’s protected class is consumers and we are willing to pay the price of some inefficiency in order to protect low consumer prices. This vision of antitrust is written into the Merger Guidelines that are in force today, which refuse to recognize an “efficiency defense” in merger cases unless the efficiency gains are large enough to guarantee that the merger will not result in higher consumer prices at all.⁴⁷ But except for this difference, which rarely shows up in litigated cases, the goals of the antitrust laws are purely economic.

One problem with admitting alternative goals is that the resulting theory becomes less robust and more susceptible to interest-group pressure “capture.” Congress is most likely to pass good legislation when goals are clearly defined and the path to getting there is relatively clear. When goals are divergent, ambiguous, or poorly articulated, Congress has much more difficulty formulating its own vision for how things should be. At that point, it becomes far more willing to listen to special interests.⁴⁸

Another problem with noneconomic approaches generally is that making them operational is virtually impossible. For example, giving serious recognition in antitrust policy to protection of small business might end up condemning every efficient practice that enables larger firms to undersell smaller ones. If pursued consistently, that policy could drive us back to the stone age, and the theory itself offers very little guidance about where to draw the line.⁴⁹

An important difference between antitrust law on the one hand and the patent and copyright laws on the other is that antitrust is passed under the commerce power, while the IP provisions are authorized by the Constitution’s Intellectual Property Clause. That fact would appear to place a

⁴⁶ See Joseph Farrell & Michael L. Katz, *The Economics of Welfare Standards in Antitrust*, COMPETITION POL’Y INT’L, Autumn 2006, at 3; John B. Kirkwood & Robert H. Lande, *The Chicago School’s Foundation Is Flawed: Antitrust Protects Consumers, Not Efficiency*, in HOW THE CHICAGO SCHOOL OVERSHOT THE MARK: THE EFFECT OF CONSERVATIVE ECONOMIC ANALYSIS ON U.S. ANTITRUST 89 (Robert Pitofsky ed., 2008).

⁴⁷ U.S. DEP’T OF JUSTICE & FED. TRADE COMM’N, HORIZONTAL MERGER GUIDELINES § 4.0, at 31 (rev. ed. 1997), <http://www.usdoj.gov/atr/public/guidelines/hmg.pdf> (Question is “whether cognizable efficiencies likely would be sufficient to reverse the merger’s potential to harm consumers in the relevant market, e.g., by preventing price increases in that market.”).

⁴⁸ See discussion *infra* text accompanying notes 70-76.

⁴⁹ See 1 PHILLIP E. AREEDA & HERBERT HOVENKAMP, ANTITRUST LAW ¶¶ 101, 111 (3d ed. 2007) (discussing alternative goals for antitrust laws and advocating a dominant economic approach).

giant thumb for the IP laws on the efficiency side of the scale. The Commerce Clause says nothing about encouraging competition or efficiency as an exclusive or even an articulated goal. It merely gives Congress the power to regulate interstate and foreign commerce. Indeed, the Commerce Clause has provided the Congressional authorization for many “fairness” provisions, such as the federal civil rights statutes.⁵⁰ In sharp contrast, the Constitution’s IP Clause expressly articulates a goal of incentivizing innovation. Indeed, the IP clause states more strongly than any other constitutional provision a goal of furthering economic growth. The purpose of patent and copyright legislation is to “promote the progress of science and useful arts.” To this end, the property rights that these provisions create must be valid for only “limited times.”⁵¹ That is, their purpose is to create incentives to innovate by giving creative people a limited period to capture enough of the benefit to incentivize their work. When that time has expired, the innovation must be given over to the public domain.

The Constitution’s purely economic authorization for IP law notwithstanding, IP writers have produced plenty of noneconomic theories that depend on natural law, common law property rights theory, Hegelian personality theory,⁵² Lockean labor theory,⁵³ and the like. In general, noneconomic theories seem to be more attractive in copyright⁵⁴ than in patent law.⁵⁵ Where

⁵⁰ *E.g.*, 42 U.S.C. § 2000e (2006) (workplace discrimination by employers).

⁵¹ U.S. CONST. art. I, § 8, cl. 8.

⁵² *E.g.*, Justin Hughes, *The Philosophy of Intellectual Property*, 77 GEO. L.J. 287, 330–65 (1988) (advocating Hegelian theory based on primacy of personal autonomy)

⁵³ *E.g.*, JEREMY WALDRON, *THE RIGHT TO PRIVATE PROPERTY* 137–252 (1988) (defending Lockean rights based on labor and individual contribution).

⁵⁴ *See, e.g.*, Stephen Breyer, *The Uneasy Case for Copyright: A Study of Copyright in Books, Photocopies, and Computer Programs*, 84 HARV. L. REV. 281 (1970) (enumerating and rejecting noneconomic goals for copyright); Wendy J. Gordon, *A Property Right in Self-Expression: Equality and Individualism in the Natural Law of Intellectual Property*, 102 YALE L.J. 1533 (1993) (advocating a modified natural law approach); Glynn S. Lunney, Jr., *Reexamining Copyright’s Incentives-Access Paradigm*, 49 VAND. L. REV. 483 (1996) (defending an allocative efficiency justification); Stewart E. Sterk, *Rhetoric and Reality in Copyright Law*, 94 MICH. L. REV. 1197 (1996) (critiquing incentives justifications); *see also* LIOR ZEMER, *THE IDEA OF AUTHORSHIP IN COPYRIGHT* (2007); Benjamin G. Damstedt, Note, *Limiting Locke: A Natural Law Justification for the Fair Use Doctrine*, 112 YALE L.J. 1179 (2003); John Tehranian, *Et Tu, Fair Use? The Triumph of Natural-Law Copyright*, 38 U.C. DAVIS L. REV. 465 (2005); Lior Zemer, *The Making of a New Copyright Lockean*, 29 HARV. J.L. & PUB. POL’Y 891 (2006) (defending a lockean contribution-of-labor theory).

⁵⁵ *See, e.g.*, Adam Moore, *Toward a Lockean Theory of Intellectual Property*, in *INTELLECTUAL PROPERTY* 81 (Adam Moore ed., 1997) (advocating Lockean, labor-driven justification); Edwin C. Hettinger, *Justifying Intellectual Property*, 18 PHIL. & PUB. AFF. 31 (1989) (similar); Hughes, *supra* note 52; Tom G. Palmer, *Are Patents and Copyrights Morally Justified? The Philosophy of Property Rights and Ideal Objects*, 13 HARV. J.L. & PUB. POL’Y. 817, 820 (1990) (personality theory); D. B. Resnik, *A Pluralistic Account of Intellectual Property*, 46 J. BUS. ETHICS 319 (2003) (advocating multi-faceted approach).

copyright is concerned, noneconomic theory seems to have the upper hand in Congress. For instance, Congress continuously and retroactively extends the term of the Copyright Act, with the result that copyright today is of effectively indefinite duration, in two different senses. First, it appears likely that Congress will keep extending and re-extending, with the result that copyrights currently in existence will never expire. And secondly, even without further extension, today's term of life of the author plus seventy years is nearly the economic equivalent of indefinite protection.⁵⁶ The result seems quite inconsistent with the mandate of the IP clause and is more akin to a natural rights or property-labor theory that creates permanent private property rights to those who have enriched society with their labor.

B. The Statutory Structure of Antitrust and the IP Laws

The differing attitudes that IP and antitrust have toward foundational issues is due in part to striking differences between the respective statutes. The antitrust laws generally condemn restraints on competition without providing specific instructions about how to achieve underlying goals. The Sherman Act, passed in 1890, gives little guidance for identifying anticompetitive practices. It condemns agreements that “restrain trade”⁵⁷ and unilateral conduct that “monopolizes,”⁵⁸ but neither of these terms is defined in the statute and the word “competition” never appears at all. This fact once led Justice Holmes to chastise his colleagues for arguing “as if maintaining competition were the expressed object of the Act.” In fact, he correctly observed, the Sherman Act “says nothing about competition.”⁵⁹

In the Clayton Act twenty-four years later, Congress said a little more, condemning in very general terms anticompetitive tying, exclusive dealing, price discrimination, and mergers. But the Clayton Act still provides very little detail, stating only that these practices are unlawful when they “may substantially lessen competition or tend to create a monopoly.”⁶⁰ Nothing in the antitrust statutes defines competition or says how it may be reduced or how much reduction is needed to

⁵⁶ See *Eldred v. Ashcroft*, 537 U.S. 186, 255–256 (2003) (Breyer, J., dissenting) (observing that under the Act, the present value of the legislatively created term was 99.8% of the value of infinitely long protection making the effective term “virtually perpetual”); Christina Bohannon, *Reclaiming Copyright*, 23 *CARDOZO ARTS & ENT. L.J.* 567 (2006) (recognizing extent of special interest capture in copyright and proposing statutory construction mechanisms for addressing it).

⁵⁷ 15 U.S.C. § 1 (2006).

⁵⁸ *Id.* § 2.

⁵⁹ *N. Sec. Co. v. United States*, 193 U.S. 197, 403 (1904) (Holmes, J., dissenting) (rejecting majority's condemnation of a railroad merger).

⁶⁰ See 15 U.S.C. § 13 (2006) (price discrimination), *id.* § 14 (tying and exclusive dealing); § 18 (mergers).

trigger a violation. Nor is there any reference to marginal-cost pricing or output-maximizing conduct, which have become the predominant baselines for measuring competition. Congress apparently did not want to get involved in articulating a specific definition of competition or in determining which practices might promote or undermine it. Rather, it enacted a few general principles derived from the common law,⁶¹ and then left it largely to the courts to determine what practices violate them.

By contrast, both the Patent Act and the Copyright Act are lengthy codes, describing in detail the kinds of rights that they create and the remedies that are available to enforce them. Patents in particular are the subject of heavy regulation, mainly through the United States Patent and Trademark Office (PTO). Why did Congress provide so little detail in the antitrust laws, simply handing that job over to the courts, but so much in the IP laws? One explanation is that the members of Congress were not economists or market experts. They did not want to tread in areas where they were poorly equipped and could do much more harm than good. So they left it to the federal courts to develop a common law of anticompetitive practices on a case-by-case basis. And that has largely been the course of antitrust policy ever since.⁶²

If that is the story for antitrust, however, it should apply even more forcefully to the IP laws. At every point in our intellectual history we have understood and been able to express the technical requirements for traditional price and output competition far more satisfactorily than the optimum requirements for furthering innovation.⁶³

Rather, differing levels of detail in the antitrust and IP provisions is most likely the result of the relative power of special interests in their drafting. The development of patent and copyright legislation has reflected the wishes of patent and copyright holders much more than the interests of consumers. In copyright, Congress has yielded increasingly to interest groups representing mainly copyright holders, particularly since passage of the 1976 Copyright Act, sometimes even permitting them to take over the process of statutory drafting.⁶⁴ Patent law has proved nearly as susceptible to

⁶¹ See 1 AREEDA & HOVENKAMP, *supra* note 49, ¶¶ 101–03 (Sherman Act derived from common law of restraints on trade).

⁶² *See id.*

⁶³ *See infra* text accompanying notes 70-74.

⁶⁴ *See* Bohannon, *supra* note 56 (extent of special interest capture in 1976 Act); Jessica D. Litman, *Copyright, Compromise, and Legislative History*, 72 CORNELL L. REV. 857, 862-79 (1987) (extent of legislative capture in 1976 Copyright Act); Neil Weinstock Netanel, *Locating Copyright Within the First Amendment Skein*, 54 STAN. L. REV. 1, 67-69 (2001) (lamenting continuous expansion of protection since 1909 resulting from Congress's reliance on interested parties to do statutory drafting); Stewart Sterk, *Rhetoric and Reality in Copyright Law*, 94 MICH. L. REV. 1197, 1244-46 (1996) ("In the period leading to the 1976 Copyright Act, Congress made it clear that industry representatives would have to hammer out a bill acceptable to all interest groups").

special interest influence.⁶⁵

To be sure, special interests have always been present in antitrust as well. With a few exceptions in the regulated industries,⁶⁶ however, they have never showed up very much in the statutory language, which has generally limited itself to open-ended statements about competition and monopoly.⁶⁷ But blaming special interest capture for the differences in the antitrust and IP laws is too superficial, for it begs the question of why the IP laws have been so much more susceptible to capture than the antitrust laws.

1. Default Positions

Two explanations seem convincing to us. The first has to do with default rules. In the United States, the underlying economics of competition policy has always been either classicism or neoclassicism. Under both of these schools of economic thought, the default rule was free and open market competition. Antitrust starts out from the position that markets generally work well and correct themselves, and government intervention on competition grounds is justified only occasionally.⁶⁸ Antitrust seeks to define the circumstances in which this occasional intervention is justified. When the antitrust laws do intervene they are mainly corrective, in the sense that they are designed to restore a generally unregulated balance that was previously upset. Even antitrust decrees that are intended to stay in place for a long time typically articulate the restoration of competitive conditions and their own expiration as a goal.⁶⁹

⁶⁵ See Robert P. Merges, *One Hundred Years of Solicitude: Intellectual Property Law, 1900-2000*, 88 CAL. L. REV. 2187, 2215–32 (2000) (Congress' susceptibility to lobbyists in passage of IP statutes); Carl Shapiro, *Patent System Reform: Economic Analysis and Critique*, 19 *Berkeley Tech. L.J.* 1017, 1021-1023 (2004) (discussing lawyers and inventors as interest groups); and see Jay P. Kesan and Andres A. Gallo, *The Political Economy of the Patent System*, 87 N.C.L.REV. 101 (2009) (influence of special interests on current patent reform efforts). On earlier patent law, see Andrew P. Morriss & Craig Allen Nard, *Institutional Choice & Interest Groups in the Development of American Patent Law: 1790-1870* (Ill. Law & Econ. Research Paper Series, Paper No. LE07-007, 2008), available at <http://ssrn.com/abstract=1262970> (role of interests groups since early Constitutional period); Craig Allen Nard, *Legal Forms and the Common Law of Patents* (Case Research Paper Series in Legal Studies, Paper No. 09-18, 2009), available at <http://ssrn.com/abstract=1396679> (advocating common law approach to control for capture); see also DAN L. BURK & MARK A. LEMLEY, *THE PATENT CRISIS AND HOW THE COURTS CAN SOLVE IT* 95–100 (2009) (tracing history of failed Congressional efforts to reform patent law and capitulation to special interests).

⁶⁶ 2A AREEDA & HOVENKAMP, *supra* note 29, ¶¶ 249–57 (3d ed. 2007).

⁶⁷ On the role of special interests in the development of antitrust legislation, see 1 *id.* ¶¶ 100–02 (3d ed. 2006). See also HERBERT HOVENKAMP, *THE ANTITRUST ENTERPRISE: PRINCIPLE AND EXECUTION*, ch. 1 (2006); Thomas J. DiLorenzo, *The Origins of Antitrust: An Interest-Group Perspective*, 5 INT'L REV. L. & ECON. 73 (1985); Herbert Hovenkamp, *Distributive Justice and the Antitrust Laws*, 51 GEO. WASH. L. REV. 1 (1982); George J. Stigler, *The Origin of the Sherman Act*, 14 J. LEG. STUD. 1 (1985).

⁶⁸ See HOVENKAMP, *ANTITRUST ENTERPRISE*, note 67 at 31-56.

⁶⁹ *E.g.*, *United States v. Microsoft Corp.* 253 F.3d 34, 103 (D.C. Cir. 2001) (en banc) (per curiam) (antitrust remedy decree “must seek to ‘unfetter a market from anticompetitive conduct,’ to ‘terminate the illegal monopoly, deny to

In sharp contrast, both the Patent Act and the Copyright Act begin with the premise—fully justified by the Constitution’s Intellectual Property Clause—that the market operating alone will not produce the optimum amount of innovation. Indeed, market failure is the starting point for IP, and it is market failure that gives rise to the need for legal entitlements. Without such entitlements, other persons could free ride by copying an innovator’s invention or copyrighted work. The result would be fewer people who are willing to make these investments and therefore fewer socially beneficial innovations. Thus, government action rather than inaction is the default rule; the only question is how much of it is desirable. That is to say, the IP laws are affirmative regulatory provisions, and regulation naturally invites special interest participation in questions about the nature and scope of government intervention.

2. Lack of Theoretical Robustness

A second reason that special interests have had more traction in IP than in antitrust has to do with the relative robustness of our competition and innovation models. The history of competition and innovation policy consistently shows more consensus about the nature of traditional competition than about the nature of innovation. The point can be pushed too far, and there is always plenty of dissent along the margins, but the principle is a strong one nonetheless. Competition requires rivalry among firms in the same market and free entry. We can debate endlessly about how many firms it takes to make a market perform competitively, or about adjustments for product differentiation, or high fixed costs or informational imbalances, but the basic outline of the requirements for competition have been well known for more than a century and claim a broad professional and policy consensus.⁷⁰

When we ask instead what government policy will encourage the optimum amount of innovation, the answers become far less determinate and go much more to the extremes. To this day, the economics of innovation has no equivalent to formulations such as the robust, broadly applied neoclassical rule that “under perfect competition price equals marginal cost.” To be sure, there are always complicated tradeoffs. For example, in antitrust we continually compare the dominant firm’s need to expand aggressively against the resulting harm to rivals, speculating about the long-run impact on consumers. But the tradeoffs in IP law are much more extreme and difficult to measure. No one knows what the optimal duration of patent or copyright protection should be, or whether there should be different periods of protection in different areas of enterprise. Some even doubt whether we need any protection at all.⁷¹ Although the patent term (twenty years from

the defendant the fruits of its statutory violation, and ensure that there remain no practices likely to result in monopolization in the future.” (citations omitted)).

⁷⁰ See, e.g., George J. Stigler, *Perfect Competition, Historically Contemplated*, 65 J. POL. ECON. 1 (1957) (tracing development of robust neoclassical model of perfect competition).

⁷¹ E.g., MICHELE BOLDRIN & DAVID K. LEVINE, *AGAINST INTELLECTUAL MONOPOLY* (2008) (two theoretical

filing of the patent application) is much shorter than the copyright term, it may still be far too long in areas such as computer technology, where inventions become obsolete in the marketplace long before their patents expire and long terms contribute to innovation-stifling thickets.. On the other hand, the patent term might be too short in areas such as pharmaceuticals, where shelf lives are very long, copying is relatively easy, and innovation costs are very high in relation to annualized payoffs.⁷² Nor are there robust models for explaining the proper scope of patent claims, the optimal amount of fair use in copyright, and the like.⁷³

About the best we can say is that the primary goal of IP policy should be to maximize net gains from innovation after all transaction costs have been paid. As such, it must balance the incentive value of exclusion against that of access to the developed technology and ideas of others.⁷⁴ This may or may not require different decisions about patent term and scope in different industries.⁷⁵ Today the consensus is fairly broad that we provide too much protection.⁷⁶

In general, the influence of special interests depends on the relative robustness of the model under which democratic policy makers work.⁷⁷ When a model is robust and obtains broad

economists arguing for abolition of IP laws; first mover advantages sufficient to facilitate most worthwhile innovation); Breyer, *supra* note 54 (finding very weak case for copyright).

⁷² See SUZANNE SCOTCHMER, *INNOVATION AND INCENTIVES* 117 (2006) (single term in all markets rewards some industries too much while rewarding others too little); Dan L. Burk & Mark A. Lemley, *Is Patent Law Technology-Specific?*, 17 *BERKELEY TECH. L.J.* 1155 (2002) (similar).

⁷³ See Andrew F. Christie & Fiona Rotstein, *Duration of Patent Protection: Does One Size Fit All?*, 3 *J. INTELL. PROP. L. & PRAC.* 402 (2008); Richard Gilbert & Carl Shapiro, *Optimal Patent Length and Breadth*, 21 *RAND J. ECON.* 106 (1990). The classic debate is WILLIAM D. NORDHAUS, *INVENTION, GROWTH, AND WELFARE: A THEORETICAL TREATMENT OF TECHNOLOGICAL CHANGE* (1969) (examining diverse factors in various industries that affect value of patent protection and producing formulas for optimal patent duration), and F. M. Scherer, *Nordhaus' Theory of Optimal Patent Life: A Geometric Reinterpretation*, 62 *AM. ECON. REV.* 422 (1972) (interpreting Nordhaus). On copyright, see Yuehong Yuan & Stephen F. Roehrig, *On the Duration of Copyright Protection for Digital Information*, in *THE INTERNET AND TELECOMMUNICATIONS POLICY* 247 (Gerald W. Brock & Gregory L. Rosston eds., 1996) (possibility of differential duration for digital and more traditional works).

⁷⁴ See, e.g., WILLIAM M. LANDES & RICHARD A. POSNER, *THE ECONOMIC STRUCTURE OF INTELLECTUAL PROPERTY LAW* 11–36 (2003) (defining optimal system in terms of “incentive-access tradeoff”).

⁷⁵ See BURK & LEMLEY, *supra* note 65; Dan L. Burk & Mark A. Lemley, *Policy Levers in Patent Law*, 89 *VA. L. REV.* 1575 (2003).

⁷⁶ See, e.g., MICHAEL HELLER, *THE GRIDLOCK ECONOMY: HOW TOO MUCH OWNERSHIP WRECKS MARKETS, STOPS INNOVATION, AND COSTS LIVES* (2008); LAWRENCE LESSIG, *FREE CULTURE: HOW BIG MEDIA USES TECHNOLOGY AND THE LAW TO LOCK DOWN CULTURE AND CENTRAL CREATIVITY* (2004). Even relative conservatives such as Landes and Posner agree. See LANDES & POSNER, *supra* note 74.

⁷⁷ See Bohannon, *supra* note 56.

consensus, the influence of special interests tends to be minimized. For example, the pricing of both groceries and retail electricity are left to the states, which means that numerous jurisdictions could intervene in either or both markets. Yet grocery prices are set by the market virtually everywhere, while retail electricity prices are regulated by government agencies virtually everywhere. Surely this is not because the electric industry has better lobbyists than the grocery trade or its special interests are better organized in every state. The explanation is a much more obvious one. Most markets have many grocery stores. For the most part, they sell commodities, and entry is easy. Accordingly, grocery markets fit the economic model for competition rather robustly. At the other end of the spectrum, retail electricity fits the economic model for a natural monopoly. We know that electricity is most efficiently delivered to houses by a single firm, which requires price regulation. Thus, although the grocery and electricity markets are very different, there is broad agreement on the appropriate economic model and policy for each market. That consensus, more than special interest capture, explains why there is regulation in one market and not in the other.

By contrast, when an area of enterprise is not well understood, law makers are more vulnerable to special interests. That is clearly the case in patent and copyright law. Either there is no factually “right answer” to problematic questions about the duration and scope of IP rights or, more likely, the answers are very complex and may vary considerably from one market to the next. Indeed, there is probably more variation in the right approach than could ever be accounted for by a single model. Lacking any consensus resembling the neoclassical vision of price competition, Congress simply listens to those who speak with the most persuasive voices. The classic public choice paradigm clearly favors IP rights holders: they are fewer in number, have individually greater stakes, and typically have interests that are much more homogenous.⁷⁸ On the other side, the users of IP rights tend to be more numerous and heterogenous. As a result, rights holders are much better organized than IP users and communicate their wishes to Congress much more effectively.⁷⁹

As a result of special interest capture, the IP laws as enacted have become much more disconnected from their purpose, which is optimum promotion of innovation.⁸⁰ At the same time, some recent judicial decisions and pending patent reform legislation show that patent law has begun its errand into the wilderness, though it is far from complete. The outlook for copyright law is bleaker. Despite voluminous scholarly commentary and powerful arguments for reform, there is little evidence today that copyright law is ready to take that journey.

⁷⁸ *Id.* at 582 (showing how copyright law bears the hallmark characteristics of special interest legislation).

⁷⁹ *See, e.g.*, DANIEL A. FARBER & PHILIP P. FRICKEY, *LAW AND PUBLIC CHOICE: A CRITICAL INTRODUCTION* (1991); DENNIS C. MUELLER, *PUBLIC CHOICE III* (2003); Herbert Hovenkamp, *Legislation, Well-Being and Public Choice*, 57 U. CHI. L. REV. 63 (1990).

⁸⁰ See discussion *infra*, parts III and IV.

III. The Beginnings of Patent Law's Errand

As others have ably observed, the patent system is currently so beset with problems that it arguably does more harm to innovation than good.⁸¹ Burk and Lemley suggest that the consensus for stronger patent protection that led to the creation of the Federal Circuit in 1982, with its broad appellate jurisdiction over cases arising under the Patent Act, has now “broken down.”⁸² Today we have gone to the opposite extreme. We have created a system that issues too many patents, many of which are only trivial improvements over prior art, while innovation would be better served if these “inventions” were left in the public domain. Second, patents have unclear boundaries, and the patent notice system does not enable innovators to know with confidence when they are treading on another’s property rights.⁸³ Third, the patent system lacks clear rules about priority of rights. As a result, developers often discover that they are infringers only after they have made a significant, irreversible investment. Paying to get out of such a predicament often costs far more than avoiding it in the first place. As a result of these three points, in many areas of enterprise, the private costs of the patent system outweigh any private gains, and market participants would be better off if the system did not exist.⁸⁴ Fourth, the patent system, like the copyright system,⁸⁵ has lacked a serious harm requirement as an element of infringement. As a result, the rights it creates are often not related in any coherent way to the underlying goal of the system, which is to encourage the optimal amount of innovation. An infringement that does not cause foreseeable harm to the patentee cannot be said to reduce the incentive to invent.

A. The Patent/Antitrust Conflict

These problems in the patent system can have severely detrimental effects on innovation and competition. Consider how patent law’s own internal problems interfere with antitrust law’s purpose of promoting economic growth through innovation and competition. Antitrust law today

⁸¹ See, e.g., JAMES BESSEN & MICHAEL J. MEURER, *PATENT FAILURE: HOW JUDGES, BUREAUCRATS, AND LAWYERS PUT INNOVATORS AT RISK* 46–72 (2008); BURK & LEMLEY, *supra* note 65.

⁸² BURK & LEMLEY, *supra* note 65, at 3. On the creation of the Federal Circuit and the controversy surrounding it, see Paul M. Janicke, *To Be or Not to Be: The Long Gestation of the U.S. Court of Appeals for the Federal Circuit (1887-1992)*, 69 *ANTITRUST L.J.* 645 (2002), and Rochelle Cooper Dreyfuss, *The Federal Circuit: A Case Study in Specialized Courts*, 64 *N.Y.U. L.REV.* 1 (1989). On the court’s impact in strengthening patent protection, see Matthew D. Henry & John L. Turner, *The Court of Appeals for the Federal Circuit’s Impact on Patent Litigation*, 35 *J. LEG. STUD.* 85 (2006).

⁸³ See BESSEN AND MEURER, *supra* note 81, at ch. 3.

⁸⁴ *Id.* On the economic cost of excessive, weak patents, see Bronwyn H. Hall & Dietmar Harhoff, *Post-Grant Reviews in the U.S. Patent System: Design Choices and Expected Impact*, 19 *BERKELEY TECH. L.J.* 989, 992-994 (2004).

⁸⁵ See Christina Bohannon, *Copyright Harm, Foreseeability, and Fair Use*, 85 *WASH. U. L. REV.* 969 (2007).

enjoys the comparative advantage of well-behaved statutory language and doctrine. While significant conflicts exist between antitrust and patent policy, most of these result from defects in the patent law system rather than in antitrust. We consider three examples: improperly brought IP infringement suits; extreme judicial deference to patent settlements; and patent continuations and holdup problems.

1. Post-Issuance Conduct: Improper Infringement Actions

The *antitrust* law of unreasonably exclusionary infringement actions is relatively clear. If an IP holder either has or threatens to have substantial market power and brings an objectively “baseless” infringement suit, then the lawsuit is deemed anticompetitive and can violate section 2 of the Sherman Act.⁸⁶ While the cases can arise under both patent and copyright, most are patent infringement suits.⁸⁷ The cases typically arise when an infringement defendant counterclaims that the patentee brought suit on an invalid or unenforceable patent, or under circumstances where no reasonable person would have anticipated success if the true facts were known. For example, the patent might be unenforceable because it was obtained by fraud or inequitable conduct before the PTO, or perhaps the patentee should have known that the defendant’s technology did not infringe.

While the case for competitive harm from improper infringement actions is strong, uncertainty and confusion in patent law often undermine these antitrust claims. All too often no one knows what the boundaries of a patent are.⁸⁸ Thus, it is often impossible to say that no reasonable person would have brought a particular infringement suit to enforce a patent.

Further, the problem is exacerbated by the fact that the Federal Circuit itself is willing to overlook fairly outrageous patent assertions. In its *Dippin’ Dots* decision, for example, the court refused to impose antitrust liability on a patentee who obtained its patent by deceiving the PTO. The court acknowledged that the patent applicant knowingly lied in a sworn statement that there had been no sales more than one year prior to the application. In fact, some 800 such sales had occurred.⁸⁹ That such sales would have rendered the product unpatentable was indisputable because of the statutory on-sale bar.⁹⁰ The Federal Circuit held that although the omission rendered

⁸⁶ See 3 PHILLIP E. AREEDA & HERBERT HOVENKAMP, ANTITRUST LAW ¶706 (3d ed. 2008).

⁸⁷ The leading patent decision is *Walker Process Equipment, Inc. v. Food Machinery & Chemical Corp.*, 382 U.S. 172 (1965). The leading copyright decision is *Professional Real Estate Investors, Inc. v. Columbia Pictures Industries, Inc.*, 508 U.S. 49 (1993). See 3 AREEDA & HOVENKAMP, *supra* note 86, ¶ 705 (3d ed. 2008).

⁸⁸ On the unacceptably high rate of reversals and Federal Circuit dissents in claim construction decisions, see John R. Thomas, *Claim Re-Construction: The Doctrine of Equivalents in the Post-Markman Era*, 9 LEWIS & CLARK L. REV. 153, 163 & n.69 (2005); Joseph Scott Miller, *Enhancing Patent Disclosure for Faithful Claim Construction*, 9 LEWIS & CLARK L. REV. 177 (2005).

⁸⁹ *Dippin’ Dots, Inc. v. Mosey*, 476 F.3d 1337, 1347 (Fed. Cir. 2007), *cert. denied*, 128 S. Ct. 391 (2007).

⁹⁰ 35 U.S.C. § 102(b) (2006) (barring patentability of product placed on the market more than one year prior to filing of

the patent unenforceable,⁹¹ it was insufficient to create antitrust liability.

There are two ways in which the court's decision makes it easier for patent holders to restrain innovation and competition. First, the court required at least two independent pieces of information showing intent to deceive the PTO. It found only one, the omitted sales themselves.⁹² The offending sales occurred in 1987, but the subsequent infringement suit was filed in 2000, some thirteen years later.⁹³ Because proving fraud would require investigation into facts that occurred many years earlier and were not generally known, the court's holding places heroic burdens on parties seeking to prove anticompetitive uses of patent infringement suits.

Second, the court's decision failed to take account of the patent holder's anticompetitive conduct in filing the infringement suit itself. The court reasoned that because antitrust penalties are far more severe than invalidating the patent, a single knowing misstatement was not sufficiently bad conduct to support an antitrust counterclaim. But the Federal Circuit overlooked the fact that there *was* in fact much more. Not only had the patentee lied to the PTO, but more than a decade later, after the PTO had closed its book on this patent and prior sales had become very difficult to trace, it filed a patent infringement claim. Thus the Federal Circuit had the two independent pieces of evidence of fraudulent conduct that it required; it simply neglected to look at the post-issuance period when the infringement action was filed. That conduct is especially important because the PTO no longer effectively regulates patentee conduct once a patent is issued. As such, the conduct is not subject to regulatory supervision and must be considered fair game under the antitrust laws.⁹⁴

When thinking about Patent Act and antitrust remedies for improper infringement activities, one must distinguish pre-issuance and post-issuance conduct. The patent application process is heavily regulated, but after a patent is issued, PTO regulation is minimal. Fundamentally, the *Walker Process* doctrine is about *post*-issuance conduct: the filing of an

the patent application).

⁹¹ See *Dippin Dots*, 476 F.3d at 1346 (“Absent explanation, the evidence of a knowing failure to disclose sales that bear all the earmarks of commercialization reasonably supports an inference that the inventor’s attorney intended to mislead the PTO. The concealment of sales information can be particularly egregious because, unlike the applicant’s failure to disclose, for example, a material patent reference, the examiner has no way of securing the information on his own.”).

⁹² *Id.* at 1346–47.

⁹³ See *In re Dippin’ Dots Patent Litig.*, 249 F. Supp. 2d 1346 (N.D. Ga. 2003) (docket entry); *cf.* *Nobelpharma AB v. Implant Innovations, Inc.*, 141 F.3d 1059 (Fed. Cir. 1998) (inventors included in patent application a reference to book published earlier that was subsequently found to be prior art precluding patentability, but the agent making the patent application deleted the reference to the book).

⁹⁴ See Herbert Hovenkamp, *Patents, Property, and Competition Policy*, J. CORP. L. (forthcoming 2009) (patent system as regulatory regime).

infringement action, or perhaps the threat to file it or threats to customers or other business relations. All of these can occur many years after a patent has issued. Antitrust should stand aside when a government agency is an active regulator, but not when economic decision making is left entirely in private hands. As a result, antitrust rightfully has a place when the anticompetitive conduct occurs subsequent to patent issuance.

The remedy of unenforceability, which the patent system itself administers against falsified applications, is designed to protect the integrity of the patent issuance process. Patents have a certain exclusionary force even if they are never enforced. For example, a firm contemplating product development in ice cream might do a patent search, discover the Dippin Dots patent, and then abandon its development efforts.⁹⁵ There may not be any correspondence whatsoever with the patentee. The result is reduced competition in a market that would have been more competitive had the invalidity of the patent been known. The antitrust laws are most likely to be invoked when another person has independently developed the technology or rightfully believed it to be in the public domain. Now the patentee brandishes the patent as a market-exclusion device. This happens during the post issuance period, when there is no oversight from the PTO.

The *Dippin Dots* case also gives us a perspective on the kinds of improper infringement actions that are most problematic. The patent applicant's sales of the patented product prior to the application are frequently undocumented and difficult to uncover. It is one thing to file a patent infringement suit on a patent that is reasonably subject to dispute on the basis of public information—for example, where there is a legal dispute about patentable subject matter⁹⁶ or claim construction.⁹⁷ Both the patentee and the infringer have the same access to relevant information—so this amounts to a “title dispute” over the meaning of the public record.

Consider the Supreme Court's 1993 *Professional Real Estate (PRE)* decision, which involved a disputed question of law.⁹⁸ The copyright holder owned the copyrights on recorded movies which PRE, who operated hotels, permitted guests to play in their rooms for a fee. The lower courts were divided on the question whether paid viewing of a movie in a hotel room constituted a “public” performance, which was a right retained by Columbia. Thus, under the broader version of the law, Columbia had a right to enjoin or collect royalties for paid viewing of

⁹⁵ On a wide variety of exclusionary conduct falling short of actual infringement suits or threats to sue, see Christopher R. Leslie, *Patents of Damocles*, 83 IND. L.J. 133 (2008).

⁹⁶ *E.g.*, *In re Bilski*, 545 F.3d 943 (Fed. Cir. 2008) (en banc), *cert. granted*, 129 S. Ct. 2735 (2009) (No. 08-964) (scope of patentable subject matter). *See also infra*, notes 175-186 and accompanying text.

⁹⁷ *E.g.*, *Kinetic Concepts, Inc. v. Blue Sky Med. Group, Inc.*, 554 F.3d 1010 (Fed. Cir. 2009) (disputed claim construction over word “wound” and other terms in method patent for wound treatment).

⁹⁸ *Prof'l Real Estate Investors, Inc. v. Columbia Pictures Indus., Inc.*, 508 U.S. 49 (1993).

its recorded films. The Supreme Court held that an infringement action was not improper simply because the lower courts disagreed with each other about the copyright holder's rights and the plaintiff chose one side rather than the other.

. But fraudulent failure to disclose prior sales does not ordinarily appear in the record, and sales made by the patentee itself thirteen years earlier are unlikely to show up at all. The thing that makes enforcement actions such as *Dippin' Dots* so dangerous is that prior sales are likely to be known only by the patent applicant and discovered by the infringement defendant only as a matter of luck. Continuing the analogy to real property, lying about prior sales is more like forgery of a signature on a deed. Such a deed fails to pass any title, but the forgery itself cannot be detected by a title search.⁹⁹

Antitrust challenges to patent infringement suits, to the threat of such suits, or to related exclusionary conduct should be analyzed as of the time the suit is brought or the threat made. At that point, the question is whether a reasonable patent owner aware of the entire record should have known that the patent was unenforceable under the circumstances. If the answer to that question is yes, then the "conduct" element of a monopolization offense has been established. Sometimes making this assessment will involve some analysis of how serious any misconduct before the PTO was, and whether the misconduct was "material," in the sense that the patent would not have issued (or would have been narrower) but for the misconduct. In all cases, however, the question is what a reasonable patent owner would have believed.

2. *Judicial Deference to IP Settlements*

Another area in which patent law enables patentees to restrain competition and innovation involves judicial deference to patent and other IP settlements. Courts have been extremely deferential to settling parties, even to the point of tolerating naked antitrust restraints. For example, in the so-called "exit payment" cases in the pharmaceutical market, courts have upheld a drug patent holder's payments to another pharmaceutical company for the other's agreement not to make and market a generic version of the drug.¹⁰⁰

⁹⁹ See, e.g., *Erier v. Creative Finance & Investments, LLC*, 349 Mont. 207, 214, 203 P.3d 744 (2009) ("A forged deed is absolutely void and wholly ineffectual to pass title, even to a subsequent innocent purchaser.").

¹⁰⁰ The Federal Circuit recently weighed in with its view of the legality of reverse payment settlements in *Hatch Waxman* patent litigation. *In re Ciprofloxacin Hydrochloride Antitrust Litig.*, 544 F.3d 1323 (Fed Cir. 2008) (rejecting Solicitor General's view that the court should examine the validity and enforceability of the patent under the circumstances), *cert. denied*, 129 S.Ct. 2828 (2009). The Sixth Circuit had previously found such payments to be per se illegal as analogous to market allocation agreements among competitors. *In re Cardizem Antitrust Litig.*, 332 F.3d 896 (6th Cir. 2003). By contrast, the Eleventh Circuit rejected per se treatment if such settlements did not restrict competition beyond the exclusionary scope of the patent themselves. *Valley Drug Co. v. Geneva Pharms.*, 344 F.3d 1294 (11th Cir. 2003); *see also In re Tamoxifen Citrate Antitrust Litig.*, 466 F.3d 187 (2d Cir. 2006) (no violation); *Schering-Plough Corp. v. FTC*, 402 F.3d 1056 (11th Cir. 2005) (no violation); MICHAEL A. CARRIER, *INNOVATION FOR THE 21ST CENTURY: HARNESSING THE POWER OF INTELLECTUAL PROPERTY AND ANTITRUST LAW* (2009)

This deferential attitude toward anticompetitive patent settlements may reflect judges' eagerness to encourage parties to resolve their disputes privately.¹⁰¹ But clearly that is not a complete answer. The real source of the problem is the complete indeterminacy that the patent system creates with respect to ownership and priority of patent rights. To illustrate, consider how antitrust law would apply to "exit payments" involving real property, where ownership rights are much clearer. Suppose Alpha owns an urban parcel of land with a working gasoline station. When Beta begins constructing a competing gasoline station across the street, Alpha files a trespass action against Beta even though Alpha has no title or possessory claim to the land on which Beta is building. The parties then "settle" the dispute by agreeing that Alpha will make "exit payments" to Beta, giving it \$1000 monthly in exchange for abandonment of its construction plans.

No court would think twice about investigating the land title records, determining that there was no property dispute here, and concluding that the whole charade was a cover for unlawful collusion. The difference in the pharmaceutical situation has nothing to do with any kind of deference to settlements generally. Rather, the reason is that the state of the "title records" in patent law is so abysmal that courts are inclined to defer to the parties' judgments about them. If patent boundaries were clearer we would almost certainly have many fewer anticompetitive settlement agreements.

As Bessen and Meurer observe,¹⁰² patent law is fraught with problems of fuzzy boundaries, overly broad claims, retroactive assertions of interest, and serious questions about validity that make the patent system a completely unreliable guide to the proper state of patent rights. These problems and uncertainties, rather than any judicial preference for settlements, drive the extreme deference toward anticompetitive patent settlements.

3. *Priority, Holdup, and Disclosure Obligations*

Finally are the problems of late patent claims and holdups, some of which the PTO has attempted to address, although thus far without much success.¹⁰³ Much of the problem of priority

(advocating presumptive illegality); C. Scott Hemphill, *An Aggregate Approach to Antitrust: Using New Data and Rulemaking to Preserve Drug Competition*, 109 COLUM. L. REV. 629 (2009); Herbert Hovenkamp, Mark D. Janis & Mark A. Lemley, *Anticompetitive Settlements of Intellectual Property Disputes*, 87 MINN. L. REV. 1719 (2003) (advocating presumptive illegality for large payment in excess of reasonably anticipated litigation costs). At this writing, a bill pending before Congress would abolish all such payments. Protecting Consumer Access to Generic Drugs Act of 2009, H.R. 1706, 111th Cong. (2009).

¹⁰¹ See 12 HERBERT HOVENKAMP, ANTITRUST LAW ¶ 2046 (2d ed. 2003).

¹⁰² BESSEN & MEURER, *supra* note 81.

¹⁰³ See discussion *infra* text accompanying notes 228-229.

and holdup in the patent system results from “late claiming,” or patent claims that are submitted to the PTO and approved subsequent to the filing of the original application. By statute, patent protection begins as of the date the application is filed,¹⁰⁴ so the result of late claiming can be the creation of property rights that are “retroactive,” in the sense that the public record first provides evidence of the right much later than the right is acknowledged for infringement purposes. Late claiming can lead to assertions of patent rights over third parties whose technological development actually preceded approval and publication of the claim that they have infringed. These parties become infringers even though they could not have had notice of the patent claims at the time of their own investment. As many as 70 percent of issued patents now claim priority to at least one previously filed patent application by virtue of backdated enforceability to the time of filing.¹⁰⁵ The impact can be catastrophic. *Prospectively*, a licensee will be willing to pay the incremental value of an invention in light of competitive alternatives, and that is the invention’s efficient value. *Retrospectively*, however, if the potential licensee has already built a factory or made another substantial commitment to the patentee’s technology without knowledge of the prior claim, the patentee can appropriate the additional value up to the cost of switching to a different technology.¹⁰⁶ This gives late claims a strategic value at the expense of innovation.

The “claims” in a patent are individual statements that set out the boundaries for determining what constitutes infringement. Using the real property deed as an analogy, the written description in a patent is a little like the general description given in the granting clause of the deed, such as “Blackacre,” or “the farm at Route 2, Box 192, Celery Center Road.” The written description was historically designed to demonstrate that the patentee was in “possession” of the invention, in the sense that he had all technical knowledge necessary to produce it and make it operational.¹⁰⁷ Just like the real property deed, however, the patent must also include a precise statement that will enable a skilled professional to know with some accuracy where the boundaries are located. In a deed this could be a metes and bounds description by reference to compass courses and distances, or it could be a reference to a survey map and to surveyors’ pins that are driven into the ground along the property lines. One who sees “O hereby grants Blackacre to X” at the top of a deed may have a rough idea of what piece of land is being granted, but only the detailed description sets forth the precise boundaries.

The written description in a patent typically describes the state of the prior art, the problems

¹⁰⁴ 35 U.S.C. § 154 (2006).

¹⁰⁵ See *PatentlyO* (July 26, 2009), available at <http://www.patentlyo.com/patent/2009/07/priority-claims-in-issued-patents.html>.

¹⁰⁶ See Mark A. Lemley & Carl Shapiro, *Patent Holdup and Royalty Stacking*, 85 TEX. L. REV. 1991, 1995 (2007).

¹⁰⁷ Timothy R. Holbrook, *Possession in Patent Law*, 59 SMU L. REV. 123 (2006) (developing diverse and idiosyncratic meanings of requirement that patent applicant be in “possession” of the invention).

that the patent addresses, and the overall nature of the invention in sufficient detail that a person skilled in the art could replicate it without too much experimentation.¹⁰⁸ In describing the entire invention, the written description will very likely refer to things that are already in the prior art, and thus not covered by this particular patent, as well as the new inventive steps that the invention includes, often without distinguishing the two. The “enablement” requirement virtually guarantees this mixture. For example, if a device contains ten components, a sufficiently enabled written description would ordinarily have to describe all ten of them, with instructions as to how they work together. It may be the case, however, that only three of the components are novel, and the inventor is claiming only these as the invention.¹⁰⁹ The claims would then identify these three components as the things that this particular patentee is claiming as novel, and for which an exclusive right is sought.¹¹⁰ As the Supreme Court has observed, the purpose of claims is to “inform the public . . . of the limits of the monopoly asserted, so that it may be known which features may be safely used or manufactured without a license and which may not.”¹¹¹

To illustrate from a well known patent validity decision, a patent on a mechanical device for permitting a sectional sofa to have parallel, independent reclining seats contained a written description of “a sectional sofa arrangement that includes a double reclining seat sofa section,” and a console between the seats that “includes a table top.” The inventor was not claiming to have invented a “sectional sofa,” however, nor a sofa with “double reclining seats,” nor the idea of a console that can be used as a table top. All of these things had been invented before by others. Rather, the invention was limited to the mechanical apparatus that enabled the individual seats on the sofa to be reclined by the use of push buttons mounted on the console.¹¹² But someone reading

¹⁰⁸ 35 U.S.C. § 112 (2006) (“The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same . . .”).

¹⁰⁹ Cf. *Evans v. Eaton*, 20 U.S. (7 Wheat.) 356, 434 (1822) (patent’s written description “mix[ed] up the new and old” such that one could not tell which part of the described machine was new and which was not; “How can that be a sufficient specification of an improvement in a machine which does not distinguish what the improvement is nor state in what it consists nor how far the invention extends?”). On the importance of this decision in creating the necessity of independent claims, see Tun-Jen Chiang, *Fixing Patent Boundaries*, ___ Mich.L.Rev. ___ (2009) (forthcoming) (George Mason Univ. Law & Econ. Research Paper Series, Paper No. 09-22, 2009), available at <http://ssrn.com/abstract=1023829>; Christopher A. Cotropia, *Patent Claim Interpretation Methodologies and Their Claim Scope Paradigms*, 47 WM. & MARY L. REV. 49, 65 (2005).

¹¹⁰ On the centrality of claims to the determination of patent scope, see Mark A. Lemley, *The Changing Meaning of Patent Claim Terms*, 104 MICH. L. REV. 101 (2005); John F. Duffy, *On Improving the Legal Process of Claim Interpretation: Administrative Alternatives*, 2 WASH.U. J.L. & POL’Y 109 (2000).

¹¹¹ *Permutit Co. v. Graver Corp.*, 284 U.S. 52, 60 (1931); see also *Markman v. Westview Instruments, Inc.*, 517 U.S. 370, 373–75 (1996) (describing centrality of patent claims).

¹¹² U.S. Patent No. 5,064,244 (filed Jan. 3, 1991).

the written description alone would not be able to tell which parts of the described invention were new, and thus patentable, and which parts were already a part of the prior art. This particular patent then went on to enumerate twenty-one separate claims, or what were intended to be statements of those elements of the described sofa that the patentee was claiming as his invention.¹¹³

The fact that the written description in a patent does not itself provide adequate notice of patent boundaries is attested by the treatment of “omnibus” claims. Having written a sufficient description, the applicant might be tempted to write a claim such as: “I claim whatever is enabled and described by the specification [including the written description], and that is not made obvious by the prior art.”¹¹⁴ But such claims are largely useless because they simply put to the searcher the duty to make her own search through the prior art to determine what is obvious and what is not.¹¹⁵ They also create the problem of too much information, because properly enabled written descriptions are often very detailed, providing a flood of technical detail, only a small portion of which constitutes the invention.

Deeds and patents differ from one another in several respects. Deeds and their descriptions are ordinarily drafted by lawyers, perhaps with the aid of a surveyor’s description. Their language is not approved by any regulatory agency, and the clerk in the recorder’s office rarely does more than ascertain that the deed is not missing obvious elements, such as the grantor’s signature. She certainly does not review the document to make sure that the description is accurate or, in most cases, even sensible. Further, she does not do a title search to determine that the description is consistent with the grantor’s record title, which is the recording system’s equivalent of “prior art.” Deed descriptions can be litigated, however, and the general rule is that a deed description that does not describe the conveyed land with sufficient accuracy that a reasonably knowledgeable person can identify its boundaries fails to pass any title whatsoever.¹¹⁶ This approach creates a very powerful

¹¹³ *Gentry Gallery, Inc. v. Berkline Corp.*, 134 F.3d 1473 (Fed. Cir. 1998) (concluding that many of the claims were not supported by the written description portion of the specification). *See also* *White v. Dunbar*, 119 U.S. 47 (1886) (claims force patentee to “define precisely what his invention is”); BURK & LEMLEY, *supra* note 65, at 11–13 (describing claiming process).

¹¹⁴ The example comes from Chiang, *supra* note 109.

¹¹⁵ *See Ex parte Fressola*, No. 93-0828, 1993 WL 311933 (B.P.A.I. May 11, 1993) (rejecting omnibus claim because it placed upon searcher the obligation to distinguish what was in the prior art from what was novel); John M. Golden, *Construing Patent Claims According to Their “Interpretative Community”: A Call for an Attorney-Plus-Artisan Perspective*, 21 HARV. J.L. & TECH. 321, 351 (2008) (describing an omnibus claim as “frank but unhelpful generality”).

¹¹⁶ *E.g.*, *In re Poteat*, 176 B.R. 734 (Bankr. D. Del. 1995) (inadequate description rendered mortgage deed invalid, and it did not matter that trustee could have determined proper description by examining land title records for prior conveyances of same parcel); *Mitchell v. Thomas*, 467 So.2d 326 (Fla. Dist. Ct. App. 1985) (description failed to locate two property lines). A narrow exception exists for clear mistakes, such as where the drafter of the deed described a different parcel than the one the parties agree upon. In such cases the court will admit parol evidence and reform the deed if necessary. *See, e.g.*, *Drake v. Hance*, 673 S.E.2d 411 (N.C.App. 2009).

default rule: the drafter of the deed has to get it right the first time. The price of not getting it right is that the deed fails. This default rule also makes the deed a remarkably accurate device for describing property ownership, and in the vast majority of cases a trained professional can look at a deed and have little difficulty in determining the property's boundaries.

The patent process works much differently. The patent applicant offers a written description with sufficient enablement that someone skilled in the art can replicate the invention, and a set of claims that are intended to be as broad as can reasonably be asserted, with ambiguities generally resolved in favor of enforceability. The patent examiner then begins a review process that tests the individual claims against both the written description and the prior art. A particular claim might fail because this particular thing was already established in the prior art, or it might fail because the claim went beyond that which the written description included—that is, because it covered something that was not shown to be a part of the proposed invention.

Many priority problems arise because the current system gives the patentee virtually unlimited bites at the apple. If one or more claims in the original patent are rejected, the patentee may file a “continuation” application which asserts new or additional claims.¹¹⁷ Further, because the effective validity date of a patent is the date that the original application was filed, such “after claiming,” or “late claiming,” can result in a patent that is effective from, say, July, 2005, even though the precise claim that ends up being infringed was drafted later and approved even later still.

The patent system does not provide adequate incentives to applicants to propose clear claims in a timely fashion.¹¹⁸ As Tun-Jen Chiang has observed, claims that are included in the original patent application are likely to be more “honest” than late-filed claims, which give the patentee the opportunity to add things not foreseen when the original application was filed or, worse yet, to hijack the inventions of third parties whose work the applicant did not anticipate.¹¹⁹ The position that patent law should strive for is one that never finds liability for infringement against technology that was deployed prior to the time that the infringement-producing claim was approved and adequately notified to the public.

Cases such as *Rambus*, recently decided by the D.C. Circuit,¹²⁰ arise because it is too easy

¹¹⁷ See Paul M. Janicke, *When Patents are Broadened Midstream: A Compromise Solution to Protect Competitors and Existing Users*, 66 U. CIN. L. REV. 7 (1997).

¹¹⁸ The lack of proper incentives to draft clear claims is suggested by the high reversal rate: the Federal Circuit reverses district court claim constructions in more than a third of appealed cases. Kimberly A. Moore, *Markman Eight Years Later: Is Claim Construction More Predictable?*, 9 LEWIS & CLARK L. REV. 231, 233, 239 (2005).

¹¹⁹ Chiang, *supra* note 109.

¹²⁰ *Rambus Inc. v. FTC*, 522 F.3d 456 (D.C. Cir. 2008), *cert. denied*, 129 S. Ct. 1318 (2009); see also *In re Dell Computer Corp.*, 121 F.T.C. 616 (1996) (computer manufacturer participated in standard-setting organization and

for patentees surreptitiously to file retroactively enforceable continuation applications on the technology of rivals. Rambus had a patent application in process that had been filed in 1990. It subsequently participated in standard-setting talks while surreptitiously filing patent continuation applications that added new claims covering the very standards that the standard-setting organization was promulgating. These claims then related back to the original 1990 application for priority purposes. Then, three years later, after firms had committed to and implemented the standardized technology, Rambus asserted its patent and demanded royalties. Rambus was then able to increase prices as well as to foreclose the standard-setting organization's consideration of alternate technologies.¹²¹

Writing patent continuation claims on the technology of others is currently lawful under patent law. As the Federal Circuit has observed:

[T]here is nothing improper, illegal or inequitable in filing a patent application for the purpose of obtaining a right to exclude a known competitor's product from the market; nor is it in any manner improper to amend or insert claims intended to cover a competitor's product the applicant's attorney has learned about during the prosecution of a patent application. Any such amendment or insertion must comply with all statutes and regulations, of course, but, if it does, its genesis in the marketplace is simply irrelevant and cannot of itself evidence deceitful intent.¹²²

Of course, the harm here does not occur from deceitful intent, but from inadequate notice of another's property rights. Until the patent system provides adequate timely notice of the scope of patent rights, the costs of the system are likely to outweigh its benefits.

Given that patents are public records giving notice of property rights,¹²³ an important patent reform would place greater responsibility on the patent applicant to communicate effective and timely notice of what he has invented. Two competing principles are at play here. One is the patentee's wish to claim everything he invented, even if it was unforeseen at the time the patent application was filed.¹²⁴ The other is the potential infringer's wish to have timely notice of

certified that it did not have IP rights on developing standards when in fact it did).

¹²¹ See Thomas F. Cotter, *Patent Holdup, Patent Remedies, and Antitrust Responses*, 34 J. CORP. L. (forthcoming 2009); see also *Qualcomm Inc. v. Broadcom Corp.*, 548 F.3d 1004 (Fed. Cir. 2008) (applying equitable estoppel so as to enjoin some enforcement of patent where patentee misrepresented to standard-setting organization that it was in process of maintaining continuation applications on covered technology), *cert. denied*, 129 S. Ct. 2182 (2009).

¹²² *Kingsdown Med. Consultants, Ltd. v. Hollister Inc.*, 863 F.2d 867, 874 (Fed. Cir. 1988); *accord Tex. Instruments Inc. v. USITC*, 871 F.2d 1054, 1065 (Fed. Cir. 1989).

¹²³ See *Boyden v. Burke*, 55 U.S. (14 How.) 575, 582 (1852) ("Patents are public records. All persons are bound to take notice of their contents, and consequently should have a right to obtain copies of them.").

¹²⁴ Defending this principle are Stephen T. Schreiner & Patrick A. Doody, *Patent Continuation Applications: How*

everything that constitutes infringement. The current system decidedly favors the former, even if it means that developers find out only after the fact that their technology infringes someone else's patent.¹²⁵ To use the perhaps overdrawn deed analogy, it would be as if the deed left some property lines ambiguous, permitting the owner to wait until a neighbor had built a house or swimming pool and then assert that the structure crossed the boundary and file a trespass action.

The Constitution's IP Clause clearly mandates that the purpose of the patent system is to incentivize invention, and the only things that create incentives are those that are anticipated at the time invention takes place. As a result, permitting a patentee to produce a claim a year after an application is filed and then assert it retroactively creates disincentives for the innocent infringer by increasing the risks of development, while doing little to incentivize the patentee.¹²⁶

As a general rule patent continuations, or late claims, are used for one of three purposes. *First*, the claim may not have been foreseen by the applicant at the time the patent application was filed. *Second*, the claim may have been foreseen by the applicant, but the applicant chose to "lie in wait" until others developed new technology so that it could write a late claim on that technology.¹²⁷ *Third*, the failure to get a claim right the first time may have been a mistake; for example, the claims may have been drafted more narrowly than intended, or they may have initially been drafted too broadly, rejected by the PTO, and then narrowed in a later submission.

Our first-cut answers to these three problems are as follows:

- Claims that were not included in the original application because the applicant did not foresee them were not part of the incentive to create the invention in the first place;¹²⁸ as a

the PTO's Proposed New Rules Undermine an Important Part of the U.S. Patent System with Hundreds of Years of History, 88 J. PAT. & TRADEMARK OFF. SOC'Y 556 (2006) (arguing for patentee's right to capture uses or rewards to invention that were not foreseen at the time of the application).

¹²⁵ See Michael Risch, *The Failure of Public Notice in Patent Prosecution*, 21 HARV. J.L. & TECH. 179 (2007).

¹²⁶ See Michael J. Meurer & Craig Allen Nard, *Invention, Refinement and Patent Claim Scope: A New Perspective on the Doctrine of Equivalents*, 93 GEO. L.J. 1947, 1993 (2005) (noting that late claiming in response to a competitor's invention necessarily involves claims that the inventor did not foresee at the time of filing the original application).

¹²⁷ Under 35 U.S.C. § 122 (2006), pending patent applications are held in secret for eighteen months subsequent to their filing date; a continuation is treated in the same way as an original application, so there is typically no notice on the record of any claim that is filed within the eighteen months prior to a patent search. This is a plausible explanation of the facts of *Rambus Inc. v. FTC*, 522 F.3d 456 (D.C. Cir. 2008) (denying antitrust liability), *cert. denied*, 129 S. Ct. 1318 (2009).

¹²⁸ Cf. Meurer & Nard, *supra* note 126, at 1998 ("[A]n inventor's incentive is not harmed much when, ex post, she is denied patent scope over technology that she did not foresee ex ante."); Bohannon, *supra* note 85 (making similar argument for copyright); *Kintner v. Atl. Comm'n Co.*, 240 F. 716, 717 (2d Cir. 1917) ("[T]he patentee is

result giving them recognition creates entitlements that were not necessary to innovation and are thus inconsistent with the constitutional mandate that the goal of the patent system is to incentivize invention. At the very least, if such claims are to be acknowledged, they should not be applied retroactively against those who made a technological choice before the claim was on record; this could be done either by making late claims effective only as of the date that they are approved and placed on record, or else by giving prior user rights to third parties who made infringing investments prior to the publication of the claim that they infringed.¹²⁹ The general goal of the system, as of any property rights system, is notice first, and infringement only later. Changing the system in this way would entirely change the patent applicant's incentive structure; it would force her to write clear and reasonable claims and get them on the public record as soon as possible.

- Strategically withheld claims operate as an affirmative disincentive to invent and are clearly contrary to constitutionally mandated patent policy; they should never be applied retroactively to technological investments made prior to the approval of the claim.¹³⁰
- The existence of innocent mistakes must be acknowledged, but the question remains of who should bear the loss. As a general rule, the common law did not penalize trespassers or other occupants by fixing mistaken property descriptions retroactively unless there was actual notice, and with good reason. A rule placing the burden of mistakes on the person who makes them creates an incentive not to make them in the first place.¹³¹ Further, it is

conclusively presumed to have known what he invented or discovered, better than did any one else, at the time he applied for a patent. This is true, even though subsequent students may perceive . . . that he disclosed methods, means, or processes having capabilities surpassing the inventor's dreams at the time attempt was made to put achievements into words.”).

¹²⁹ See Carl Shapiro, *Prior User Rights*, 96 AM. ECON. REV. (PAPERS & PROC.) 92 (2006); FED. TRADE COMM'N, TO PROMOTE INNOVATION: THE PROPER BALANCE OF COMPETITION AND PATENT LAW AND POLICY 16 (2003), available at <http://www.ftc.gov/os/2003/10/innovationrpt.pdf>.

¹³⁰ See Robert P. Merges, *Software and Patent Scope: A Report from the Middle Innings*, 85 TEX. L. REV. 1627, 1653–54 (2007) (referring to strategic delay in filing claims and then writing them on the technology of third parties as “misappropriation by amendment”).

¹³¹ Another reason that late claiming should not be used to fix mistakes is that the doctrine of equivalents can help the patentee out by adding breadth to claims that were initially written too narrowly. See also Douglas Lichtman, *Substitutes for the Doctrine of Equivalents: A Response to Meurer and Nard*, 93 GEO. L.J. 2013 (2005) (noting extent to which doctrine of equivalents and patent continuations act as substitutes for one another—the former tends to read a broader range of coverage into existing claims by judicial decision while continuations tend to permit patentees to write broader claims for themselves after the fact); Meurer & Nard, *supra* note 126; Joshua D. Sarnoff, *The Doctrine of Equivalents and Claiming the Future After Festo*, 14 FED. CIR. B.J. 403 (2005); Joshua D. Sarnoff, *The Historic and Modern Doctrine of Equivalents and Claiming the Future, Part I (1790-1870)*, 87 J. PAT. & TRADEMARK OFF. SOC'Y 371 (2005); Joshua D. Sarnoff, *The Historic and Modern Doctrines of Equivalents and Claiming the Future: Part II (1870-1952)*, 87 J. PAT. & TRADEMARK OFF. SOC'Y 441 (2005).

much less costly for someone to draft claims clearly early on—particularly if they are properly limited to that which is anticipated—than to put third party investors at risk of infringement for something that they could not reasonably discover.¹³²

B. The Judicial Road to Patent Reform

Despite these severe problems, there is reason to believe that patent law has begun its errand toward reform. Although the late claiming problem still awaits a satisfactory solution, several judicial decisions have modified patent doctrine in other areas in an attempt to limit patentability to truly “nonobvious” inventions, provide clearer boundaries and notice of patented inventions, and limit the scope of injunctive relief.

1. *Nonobviousness and KSR*

The United States Patent Act requires a sufficient difference between a claimed invention and the existing “prior art” that a person with “ordinary skill” in that art would not regard the claimed innovation as obvious.¹³³ The statute does a number of things. *First*, it refers to “prior art” but does not specify its content. The case law indicates that “prior art” includes but is not limited to previously issued patents and many patent applications, as well as other available knowledge that pertains to the subject matter of the patent.¹³⁴ *Second*, the statute refers to an imaginary person with “ordinary skill” in this prior art, which means someone who is a trained professional in the subject area of the patent, but presumably not a genius. *Third*, the statute relates this knowledge to the process by which a patent may be “obtained,” thus indicating that it is the responsibility of the patent examiner as well as the applicant to examine the prior art to determine whether the claimed innovation takes a sufficient “inventive step.”¹³⁵ The latter term

¹³² F. Scott Kieff, *The Case for Registering Patents and the Law and Economics of Present Patent-Obtaining Rules*, 45 B.C. L. REV. 55, 99 (2003) (the patentee is in the least cost position to avoid ambiguities in the notice and search process).

¹³³ See 35 U.S.C. § 103(a) (“A patent may not be obtained ... if the differences between the subject matter sought 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.”). The requirement was not always an explicit part of U.S. patent policy, although some version of it was certainly implied. See John F. Duffy, *Inventing Invention: A Case Study of Legal Innovation*, 86 TEX. L. REV. 1 (2007).

¹³⁴ See, e.g., *Atlana Pharma AG v. Teva Pharms. USA, Inc.*, 566 F.3d 999, 1004 (Fed. Cir. 2009) (scholarly articles as prior art); *Kinetic Concepts, Inc. v. Blue Sky Med. Group, Inc.*, 554 F.3d 1010 (Fed. Cir. 2009) (articles, books, expert testimony on historical uses as prior art).

¹³⁵ *Graham v. John Deere Co.*, 383 U.S. 1, 18 (1966) (“the primary responsibility for sifting out unpatentable material lies in the Patent Office”).

is also used in the patent statutes of some foreign jurisdictions.¹³⁶ The basic idea is intuitively clear: we want to know whether an average person working in this particular discipline could readily have anticipated the subject matter of the claimed patent. If so, it is not patentable.

The nonobviousness provision and its interpretation raise some issues that go directly to the question whether we grant too many patents. Should we patent only works of great creative genius, such as steamboats, telegraphs, light bulbs, or Prozac? Or is it enough if the patentee simply comes up with a clever little gadget that does something a little differently than it had been done before? Closely related is the question of “hindsight bias” that has troubled the courts. *Ex post*, we might look at some new thing, such as a garbage bag with a pumpkin’s face printed on it, and say “that’s obvious,” or “I could have thought of that.” But the fact is that both trash bags and artistically drawn Jack-O-Lantern faces had been around for decades, and apparently no one had thought about putting them together on a commercial product.¹³⁷

More generally, what is the appropriate domain of “prior art?” Does it refer only to previously issued patents and patent processes and technical journals? Or does it also include things that fall into the category of common sense or perhaps even historical practice? This question is of enormous practical importance because patent examiners are overworked officials who can realistically dedicate only so many hours to a given patent application and can readily look in only a limited number of places.¹³⁸ The PTO maintains an enormous database of prior art, focusing mainly on patents and patent applications as well as some technical materials.¹³⁹ Private databases are available as well.¹⁴⁰ The examiner will consult the database, but patent applicants are also required to disclose known prior art. A failure to disclose could result in patent unenforceability.¹⁴¹

¹³⁶ *E.g.*, EUROPEAN PATENT CONVENTION art. 52 (13th ed. 2007), available at <http://www.epo.org/patents/law/legal-texts/epc.html>.

¹³⁷ *In re Dembiczak*, 175 F.3d 994 (Fed. Cir. 1999) (pre-*KSR*; finding patent on garbage bags bearing pumpkin face to be sufficiently nonobvious). For a critique, see Gregory N. Mandel, *Patently Non-Obvious: Empirical Demonstration that the Hindsight Bias Renders Patent Decisions Irrational*, 67 OHIO ST. L.J. 1391 (2006).

¹³⁸ See Andrew Chin, *Artful Prior Art and the Quality of DNA Patents*, 57 ALA. L. REV. 975 (2006) (tendency of overworked examiners to concentrate on previously issued patents and patent applications).

¹³⁹ See U.S. Patent and Trademark Office, Patent Full-Text and Full-Page Image Databases, <http://patft.uspto.gov> (last visited July 30, 2009).

¹⁴⁰ See, *e.g.*, ip.com, Searching the Prior Art Database, <http://www.ip.com/prior-art-database/searching.jsp> (last visited July 30, 2009).

¹⁴¹ *E.g.*, *Nobelpharma AB v. Implant Innovations, Inc.*, 141 F.3d 1059 (Fed. Cir. 1998) (failure to include reference to previously published book that was known to the drafter of the patent application).

The usefulness of searches for prior art is greatly driven by the nature of the market. Often when new technologies emerge, participants “transport” ideas from old settings into the new ones. *Amazon.com*’s famous patent on “one click shopping” is a good example.¹⁴² The main ingredients in one click shopping are a customer who can store log-in information, a valid credit card number, and a shipping address. Then, presumably with a single mouse click, the customer can select an item and *Amazon.com* automatically charges the item to the stored credit card and ships it to the stored shipping address. For purposes of argument, we can concede that *Amazon.com* was the first to come up with this innovation in internet shopping. But what about more traditional markets? Since time immemorial stores have been permitting regular customers to open accounts. The stores maintained names, delivery addresses, and perhaps some credit information on index cards. So when your pre-World-War II grandfather called the store and asked for a sack of flour to be delivered, the shopkeeper would load it into the pickup, charge the account, and deliver the flour to the address on the account slip. At the end of the month, your grandfather would get the bill. Should *Amazon.com* be entitled to collect a royalty from the thousands of internet vendors who simply adapted to the internet an age-old way of making shopping easier? The courts properly doubted it.¹⁴³ The example shows that the reason that patents such as one-click shopping are granted in the first place is because relevant prior art is construed too narrowly.¹⁴⁴

In *KSR International Co. v. Teleflex Inc.*,¹⁴⁵ the Supreme Court tightened up the standard for demonstrating nonobviousness of patent claims. There, the Court rejected the Federal Circuit’s rigid application of the teaching-suggestion-motivation (TSM) test, although not necessarily the test itself. Under the Federal Circuit’s version of the test, a patent claim was obvious only if there was a specific “teaching” or “suggestion” in the prior art regarding the combination of previously known elements into the new element that constitutes the subject of the patent claim at issue. The Supreme Court unanimously held that a patent claim is obvious if

¹⁴² Method and System for Placing a Purchase Order Via a Communications Network, U.S. Patent No. 5,960,411 (filed Sept. 12, 1997).

¹⁴³ *Amazon.com, Inc. v. BarnesandNoble.com, Inc.*, 239 F.3d 1343, 1363 (Fed. Cir. 2001) (denying preliminary injunction after finding likely obviousness, given the prior art).

¹⁴⁴ See Susan Walmsley Graf, Comment, *Improving Patent Quality Through Identification of Relevant Prior Art: Approaches to Increase Information Flow to the Patent Office*, 11 LEWIS & CLARK L. REV. 495, 500–01 (2007) (noting the problem, particularly with reference to software patents, where examiners tend to focus prior art searches on earlier patents even though the history of software patents is short); see also Robert P. Merges, *As Many as Six Impossible Patents Before Breakfast: Property Rights for Business Concepts and Patent System Reform*, 14 BERKELEY TECH. L.J. 577 (1999) (noting problem in new technologies when prior art searches focus too much on previously issued patents and patent applications).

¹⁴⁵ *KSR Int’l Co. v. Teleflex Inc.*, 550 U.S. 398 (2007).

“a person having ordinary skill in the art” (PHOSITA) *could have* anticipated the innovation offered by the patent. This was in contrast to the Federal Circuit’s application of the TSM test, in which it found a claim obvious only if someone had *already* anticipated it, or at least the need for it.

The *KSR* rule seems eminently sensible. When a person having ordinary skill in the art can foresee the benefit in doing what the patent claims, the claimed improvement seems obvious to try. The principal value of the TSM test is for *rejecting* claims, because the fact that the TSM is already in the prior art indicates obviousness. But the TSM test cannot be dispositive for accepting claims, because some things that are not taught or suggested in the prior art might still be foreseeable to the PHOSITA. A forward-looking test is particularly important in young and fast-developing markets, such as high technology and the internet. In those markets, the amount of conventional prior art is much less voluminous than in older well-established markets, and the pace of innovation is fast enough that many developments will be anticipated before they are documented in the prior art.¹⁴⁶

2. *Tactile Subject Matter and Overly Abstract Patent Claims*

The enormous case law and literature on patent subject matter covers such things as the patentability of living organisms, genetic mutations, abstract ideas, the laws of nature, or business methods.¹⁴⁷ One problem particularly important to competition policy and innovation is excessive abstraction, which allows patent claims to spill over into unanticipated areas, permitting patentees to monopolize things that they really did not invent and perhaps did not even foresee.¹⁴⁸ When the claims in issued patents are too abstract, later innovators cannot determine whether their work is covered by another’s patent. Nothing suppresses irreversible investment in innovation more than a substantial fear that once the innovation is developed, a court will find that it in fact belongs to someone else. The vaguely defined category of “business method” patents is particularly vulnerable to excessive abstraction. As a result, they have produced so much litigation that the net welfare effects are very likely negative. This places such patents on a collision course with the Constitution’s IP clause, which does not authorize patents

¹⁴⁶ See *Merges*, *supra* note 144; accord Tun-Jen Chiang, *A Cost-Benefit Approach to Patent Obviousness*, 82 ST. JOHN’S L. REV. 39 (2008).

¹⁴⁷ See, e.g., *Diamond v. Chakrabarty*, 447 U.S. 303 (1980) (patentability of synthesized bacteria not found in nature); *AT&T Corp. v. Excel Commc’ns, Inc.*, 172 F.3d 1352 (Fed. Cir. 1999) (patentability of process used to establish the value of a number used in customer billing); *State St. Bank & Trust Co. v. Signature Fin. Group, Inc.*, 149 F.3d 1368 (Fed. Cir. 1998) (patentability of business method).

¹⁴⁸ Many of the same issues arose in the early development of process patents. See 1 DONALD S. CHISUM, CHISUM ON PATENTS § 1.03 (2006).

that retard rather than promote innovation.¹⁴⁹

To illustrate, there is nothing inherently wrong about patenting methods that improve the way business is carried out, and such patents have been around for a long time. What has become particularly troublesome, however, is a high level of abstraction in these patents, coupled with the applicant's natural tendency to draft patent claims that extend to the largest possible boundaries. Overly abstract claims can exclude from domains far larger than any sensible view of the scope of the invention.

A case in point is the infamous *eData* patent, which at its core covered sale of electronic files such as music downloaded at the point of sale in shopping mall kiosks. However, the patent claims included coverage of all sales of information that was loaded into a physical object at the point of sale if the data were transported from a remote source.¹⁵⁰ The patentee later filed successful infringement suits against many sellers who engaged in such services as selling software, music, eBooks, and other digital technologies over the internet.¹⁵¹ Bear in mind that the *eData* inventor did not invent any part of the *technology* for transmitting a computer file over a communication line and then recording it onto a tape or other storage device. That technology was already in place and included such things as hard drives, software and file formats, telecommunications and modems, and cassette recorders, to name a few. Rather, the patent claimed the *idea* of telecommunicating electronic files and charging money for them at the receiving end. The patent's description and drawings are simple, empty boxes bearing titles such as "information file unit" and "reproduction unit." These referred generically to the full range of devices including hard drives or CD-burners that a file transfer system might contain.¹⁵² Under this patent, if a store (1) electronically downloaded digital songs from a remote supply source such as *iTunes*, (2) copied them onto CDs and (3) sold them to customers on site, it would be committing infringement, even though all three of these activities were accomplished with technologies that either existed prior to the time of the patent application or were invented by

¹⁴⁹ See *Graham v. John Deere Co.* 383 U.S. 1, 5–6 (1966) (“[The Constitution’s IP] clause is both a grant of power and a limitation.”); see also BESSEN & MEURER, *supra* note 81, at 25–27, 151–61, 187–88, 246–57; Malla Pollack, *The Multiple Unconstitutionality of Business Method Patents: Common Sense, Congressional Consideration, and Constitutional History*, 28 RUTGERS COMPUTER & TECH. L.J. 61, 90 (2002) (framers objected to business method patents because they believed that monopolies on trade itself were odious); Robert Patrick Merges & Glenn Harlan Reynolds, *The Proper Scope of the Copyright and Patent Power*, 37 HARV. J. ON LEG. 45, 52–53 (2000) (arguing against extensions of patent scope).

¹⁵⁰ System for Reproducing Information in Material Objects at a Point of Sale Location, U.S. Patent No. 4,528,643 (filed Jan. 10, 1983).

¹⁵¹ See *Interactive Gift Express, Inc. v. Compuserve Inc.*, 256 F.3d 1323 (Fed. Cir. 2001) (construing claims broadly). On the *eData* saga, see BESSEN & MEURER, *supra* note 81, at 1–25, 194–99.

¹⁵² See the patent, note 150.

someone else later.¹⁵³

One can approach the abstraction problem that *eData* raises in different ways. One is to resort to a system of “central claiming” as was once common in the judicial interpretation of patent claims. Under that system, the courts tried to assess patent scope by taking a less technical but more common sense look at the core of an invention, focusing more on the written description rather than considering the outermost reaches of its claims.¹⁵⁴ Another, indicated by the Federal Circuit’s *Bilski* decision, is to tie patent eligibility under §101 of the Patent Act¹⁵⁵ more securely to tactile subject matter. *Bilski*’s “machine/transformation” test queries whether a particular method or process for which a patent is sought (1) is tied to a particular machine or apparatus; or (2) transforms an article into a different state or thing.¹⁵⁶

The *Bilski* approach to eligible subject matter for process patents is hardly new. Prior to the Civil War, the Supreme Court observed how overreaching can occur when claims become too abstract. In 1853, the Court considered the validity of a reissued patent for the Morse telegraph. Samuel F. B. Morse’s patent contained eight claims. Five of them covered the functional aspects of the mechanical transmitter and receiver that Morse claimed were unique, and that distinguished his invention from those of numerous rivals.¹⁵⁷ Another claim was on the set of dots and dashes that came to be known as Morse code and another was on the use of the Morse code in combination with the hardware described in the earlier claims. In addition, however, Morse included this eighth claim:

‘I do not propose to limit myself to the specific machinery, or parts of machinery, described in the foregoing specifications and claims; the essence of my invention being the use of the motive power of the electric or galvanic current, which I call electro-magnetism, however developed, for making or printing intelligible characters, letters, or signs, at any distances, being a new application of that

¹⁵³ See *Interactive Gift Exp., Inc. v. Compuserve, Inc.*, 256 F.3d 1323 (Fed. Cir. 2001) (upholding *eData* patent against technologies that were not yet developed when patent was written, and applying it to firms that downloaded remote files, including books, to retail store and burned them on to consumer-supplied media).

¹⁵⁴ See Dan L. Burk & Mark A. Lemley, *Fence Posts or Sign Posts: Rethinking Patent Claim Construction* (2009), available at <http://ssrn.com/abstract=1358460>.

¹⁵⁵ 35 U.S.C. § 101 (2006) (permitting a patent to one who “invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof”).

¹⁵⁶ *In re Bilski*, 545 F.3d 943 (Fed. Cir. 2008) (en banc), cert. granted, 129 S. Ct. 2735 (2009) (No. 08-964).

¹⁵⁷ On the rivalry in Europe and America to develop the first commercially feasible telegraph, see GEORGE B. PRESCOTT, *HISTORY, THEORY, AND PRACTICE OF THE ELECTRIC TELEGRAPH* 50–80, 125–35 (1860); KENNETH SILVERMAN, *LIGHTNING MAN: THE ACCURSED LIFE OF SAMUEL F. B. MORSE* (2004).

power, of which I claim to be the first inventor or discoverer.¹⁵⁸

The Supreme Court approved Morse's claims on the device described in the patent specification and the code. However, Chief Justice Taney's opinion for the Court rejected the eighth claim. Taney observed that during the period leading up to Morse's patent application many scientists had been studying a variety of ways to use electricity plus magnets to send long-distance communications. Morse devised a particular system that permitted the signals to be sent long distances, rather than quickly fading in a few hundred feet, although some European scientists had similar successes at about the same time. Morse's eighth claim covered not only what he had actually invented, but the entire principle of using electromagnetic impulses to communicate over long distances. Just like the patent approved by the Federal Circuit in the *eData* case a century and a half later,¹⁵⁹ Morse was trying to commandeer all future technologies for accomplishing something. As Chief Justice Taney complained:

For aught that we now know some future inventor, in the onward march of science, may discover a mode of writing or printing at a distance by means of the electric or galvanic current, without using any part of the process or combination set forth in the plaintiff's specification. His invention may be less complicated—less liable to get out of order—less expensive in construction, and in its operation. But yet if it is covered by this patent the inventor could not use it, nor the public have the benefit of it without the permission of this patentee.¹⁶⁰

Taney's observations were prescient, because Morse's telegraph system was in fact quite primitive. Tremendous progress in telegraph technology was on the horizon, including faster keys, receivers that produced sounds rather than dots and dashes on strips of paper, and eventually the development of teletype devices that printed letters of the alphabet at the receiving end. If accepted, Morse's eighth claim would likely have delayed many of these inventions.¹⁶¹ Taney's observation also indicates that excessive abstraction is a *competition* problem. Overly broad claims eliminate rivalry because the patent covers not only the technology that the patentee actually invented, but other technologies that might have entered the market had the patent not squelched them.

¹⁵⁸ *O'Reilly v. Morse*, 56 U.S. (15 How.) 62, 86 (1854) (quoting patent claim eight).

¹⁵⁹ See discussion *supra*, text at notes 150-153.

¹⁶⁰ *O'Reilly*, 56 U.S. at 113.

¹⁶¹ See ANTON A. HUURDEMAN, *THE WORLDWIDE HISTORY OF TELECOMMUNICATIONS* 300–20 (2003) (detailing technological improvements in the telegraph subsequent to Morse's invention); ROBERT LUTHER THOMPSON, *WIRING A CONTINENT: THE HISTORY OF THE TELEGRAPH INDUSTRY IN THE UNITED STATES, 1832-1866* (1947) (similar; focusing on United States).

Early on, courts invalidated patents for their lack of tactile engagement. As the Supreme Court observed in 1868: “But if the subject-matter be neither a machine nor a manufacture, nor a composition of matter, then . . . it *must* be an art, for there can be no valid patent except it be for *a thing made*, or for the art or *process of making a thing*.”¹⁶² Early twentieth century courts began upholding what might be characterized as business method patents, *provided* that the subject matter of the patent was a tangible good or a process for changing a good in some useful way. In 1911, a federal appeals court sustained a patent on a scrip book containing coupons that could be exchanged for railway tickets on a dollar rather than mile basis, thus permitting them to be interchanged among different railroads that had differing fare structures. The court concluded:

Nor do we think that this patented concept is nothing more than a business method. Its use is a part of a business method. The ticket patented is not a method at all, but a physical tangible facility, without which the method would have been impracticable, and with which it is practicable.¹⁶³

Another court approved a patent on a perforated railroad ticket that could be torn in specific places to show which part of a route a passenger had already traveled and which parts of the ticket were still useable.¹⁶⁴ The court emphasized that the patent “describe[d] a distinctive physical structure.”¹⁶⁵ By contrast, if the method consisted only of a new use of an established device, the courts were hostile. A 1903 decision refused to permit a patent on the use of an ordinary business ledger to create entries that would track railroad cargo to ensure that it was not sent to the wrong destination.¹⁶⁶ Other courts denied patentability to bookkeeping systems that simply employed pre-existing ledger books to monitor restaurant waiters’ receipts or to track bad debts.¹⁶⁷ The physical objects affected by these patents were all well established prior art. At

¹⁶² *Jacobs v. Baker*, 74 U.S. (7 Wall.) 295, 297 (1869) (quoting GEORGE TICKNOR CURTIS, A TREATISE ON THE LAW OF PATENTS FOR USEFUL INVENTIONS (3d ed. 1866)).

¹⁶³ *Rand, McNally & Co. v. Exchange Scrip-Book Co.*, 187 F. 984, 986 (7th Cir. 1911).

¹⁶⁴ *Cincinnati Traction Co. v. Pope*, 210 F. 443 (6th Cir. 1913).

¹⁶⁵ *Id.* at 446.

¹⁶⁶ *Hocke v. N.Y. Cent. & H. R. R. Co.*, 122 F. 467 (2d Cir. 1903).

¹⁶⁷ *Hotel Sec. Checking Co. v. Lorraine Co.*, 160 F. 467 (2d Cir. 1908) (method of monitoring waiters’ order slips in order to prevent fraud unpatentable); *United States Credit Sys. Co. v. Am. Credit Indemnity Co.*, 53 F. 818, 819 (C.C.S.D.N.Y. 1893) (bookkeeping system designed to track bad debts not patentable); see Alan L. Durham, “*Useful Arts*” in the *Information Age*, 1999 B.Y.U. L. REV. 1419 (tracing how early “business method” patents were in fact approved if they involved the production or transformation of a physical object); see also P. J. Federico, *Origin and Early History of Patents*, 11 J. PAT. OFF. SOC’Y 292 (1929).

most, the patentee was contributing a different way of using them.

In sum, the courts have emphasized since the nineteenth century that patents must relate in some formative way to a physical object that is described in the written description of a patent application. As the *Morse* telegraph decision recognized, when the subject matter of patents becomes more abstract, patent claims have a tendency to reach out into unanticipated places. Morse had been pretty honest about this—his patent claim reached to all technologies “however developed,”¹⁶⁸ indicating that he was asserting control over devices that he himself could not foresee. The chill on innovation could hardly be more severe—those knowing of the patent would know that any device they made for improving electro-magnetic long distance communications would fall under Morse’s claim and constitute infringement, no matter how novel and important it was in fact.¹⁶⁹

Patents that do not create or transform specific physical objects rarely have clear boundaries.¹⁷⁰ Recognizing this, the courts have imposed a number of limitations. For example, the Supreme Court has said that you cannot patent an abstract idea, or it has compared purely mental patents to laws of nature, which cannot be patented.¹⁷¹ The rise of the computer and later the internet have led to an explosion of process patents in data processing, financial services and banking, transaction facilitation, risk determination, tax planning, educational materials provided on the internet, and even dating strategies.¹⁷² What many of these patents have in common is that their principal useful content consists of formulas or algorithms that can potentially have a wide variety of uses. Patenting a formula itself, in contrast to a specific and concrete application of a formula, can come dangerously close to patenting either a law of nature or an abstract idea, which patent policy forbids. The Supreme Court made that distinction seventy years ago in acknowledging a patent for a radio antenna whose length and attitude were based on a well known mathematical formula (the “Abraham’s formula”) relating wave length to optimal antenna length and position. The Court concluded that “[w]hile a scientific truth, or the mathematical

¹⁶⁸ See *supra* text accompanying note 158.

¹⁶⁹ Cf. F. Russell Denton, *Plumb Lines Instead of a Wrecking Ball: A Model for Recalibrating Patent Scope*, 16 J. INTEL. PROP. L. 1 (2008) (acknowledging problems caused by excessive abstraction, but also observing that writing concrete claims is extremely tedious work).

¹⁷⁰ See BESSEN & MEURER, *supra* note 81, at 186–214 (focusing mainly on software patents).

¹⁷¹ See *Diamond v. Diehr*, 450 U.S. 175, 185 (1981) (excluding from patentable subject matter “laws of nature, natural phenomena, [or] abstract ideas”); *Gottschalk v. Benson*, 409 U.S. 63, 67 (1972) (“Phenomena of nature, though just discovered, mental processes, and abstract intellectual concepts are not patentable, as they are the basic tools of scientific and technological work.”).

¹⁷² See, e.g., Internet-Based Education Support System and Methods, U.S. Patent No. 6,988,138 (filed June 30, 2000); Method of Designating Dating Status, U.S. Patent No. 7,255,277 (filed Apr. 6, 2006).

expression of it, is not a patentable invention, *a novel and useful structure created with the aid of knowledge of scientific truth* may be.”¹⁷³ The Court then limited the scope of the patent to specific structures that applied the mathematical formula.¹⁷⁴

The patent application in *Bilski* involves a series of hypothetical transactions between buyers and sellers of a particular commodity who had different risk profiles.¹⁷⁵ Some tended to profit when the value of the commodity increased while others profited when it decreased. In sum, *Bilski* was attempting to patent a very general form of risk “hedging.” Suppose that a coal burning electric utility must purchase large quantities of coal and sell its power at regulated rates. The utility suffers if the price of coal goes up, but coal mining companies benefit. As a result, an optimal investment strategy for a utility might be to purchase some stock in coal mining companies, which means that its operating losses from high coal prices would be offset by gains in the stock’s value. Hedging is nothing more than a way of managing risks, and the financial literature has been filled with such strategies for decades.¹⁷⁶ In rejecting the claim, the examiner concluded that “the invention is not implemented on a specific apparatus and merely manipulates [an] abstract idea and solves a purely mathematical problem without any limitation to a practical application.”¹⁷⁷

The important factors about the *Bilski* patent are that (1) it was not a patent on a specific apparatus, or machine; and (2) it was not for a process that would make (or improve) a specific apparatus or machine. Rather, the patent was on a set of mental steps leading to a formula for computing the optimal balance of investment for minimizing risk. The literature on hedging is highly mathematical, and different hedging strategies produce different formulas and can be applied in a variety of settings, from commodity pricing, energy management, foreign currency exchange rates, construction risks, and insurance, to name a few.¹⁷⁸ Even the process described in the *Bilski* patent application of initiating hypothetical transactions in markets with differing risk profiles is well established in the literature.¹⁷⁹

¹⁷³ *Mackay Radio & Tel. Co. v. Radio Corp. of Am.*, 306 U.S. 86, 93-94 (1939) (emphasis added).

¹⁷⁴ *Id.* at 98-99.

¹⁷⁵ *In re Bilski*, 545 F.3d 943, 954 (Fed. Cir. 2008) (en banc), *cert. granted*, 129 S. Ct. 2735 (2009) (No. 08-964).

¹⁷⁶ *See, e.g.*, CAROL ALEXANDER, *MARKET RISK ANALYSIS: PRICING, HEDGING AND TRADING FINANCIAL INSTRUMENTS* (2008); NIALL FERGUSON, *THE ASCENT OF MONEY: THE FINANCIAL HISTORY OF THE WORLD* (2008).

¹⁷⁷ *See Ex parte Bilski*, 2006 WL 5738364 (B.P.A.I. Sept. 26, 2006) (quoting patent examiner) (internal quotation marks omitted) (alteration in original).

¹⁷⁸ *See* PATRICK CUSATIS & MARTIN THOMAS, *HEDGING INSTRUMENTS AND RISK MANAGEMENT* (McGraw-Hill 2005); ROBERT W. KOLB, *UNDERSTANDING FUTURES MARKETS*, 135–50, 184–91 (5th ed. 1997).

¹⁷⁹ *See* R. H. Snape & B. S. Yamey, *Test of the Effectiveness of Hedging*, 73 J. POL. ECON. 540 (1965) (proposing

The specification in the *Bilski* patent described a risk-management process for energy markets where prices fluctuated mainly due to weather. But the first claim in the patent referred generally to “commodities,” indicating that the patent would later be asserted in any market subject to risk differentials. The *Bilski* patent, just as the one in *eData*, is a little like a deed to Camelot—it cannot be located and its boundaries are impossible to identify. These patents spill over from things that are arguably novel to processes that have been well known for centuries. An insurance company trying to diversify its risks or a farmer selling grain futures to hedge against a low market price for her crop might discover that she owes *Bilski* a royalty.

The Federal Circuit responded with its double-branched machine/transformation” test:

[A]n applicant may show that a process claim satisfies §101 either by showing that his claim is tied to a particular machine, or by showing that his claim transforms an article. Certain considerations are applicable to analysis under either branch. First, . . . the use of a specific machine or transformation of an article must impose meaningful limits on the claim's scope to impart patent-eligibility. . . . Second, the involvement of the machine or transformation in the claimed process must not merely be insignificant extra-solution activity.¹⁸⁰

The *Bilski* problem would be much less substantial if the patent laws contained a right of independent discovery, as the copyright and trade secret laws do.¹⁸¹ One cannot be liable for copyright or trade secret infringement unless they actually had access to the owner’s writing, product, or process and copied from it in some way. Any property system that makes infringers strictly liable for trespasses must include a mechanism for giving investors adequate and timely notice of their property interests. As the Supreme Court observed already in the 1960s, until a “process claim has been reduced to production of a product shown to be useful, the metes and bounds of that monopoly are not capable of precise delineation. It may engross a vast, unknown, and perhaps unknowable area.”¹⁸²

This problem is exacerbated by lingering effects of the doctrine of equivalents, which gives patentees infringement claims for uses that do not literally infringe the claims of the patent.

the use of series of hypothetical transactions to test outcomes from hedging strategies).

¹⁸⁰ *Bilski*, 545 F.3d at 961–62 (citations omitted).

¹⁸¹ See Samson Vermont, *Independent Invention as a Defense to Patent Infringement*, 105 MICH. L. REV. 475 (2006); see also Christopher A. Cotropia & Mark A. Lemley, Copying in Patent Law (2009), available at <http://ssrn.com/abstract=1270160> (suggesting that most alleged patent infringers are not copyists).

¹⁸² *Brenner v. Manson*, 383 U.S. 519, 534 (1966).

Patent infringement claims under the doctrine of equivalents are permitted even when the innovator developed its innovation independently. The doctrine makes intuitive sense when applied to someone who intentionally invents around a known patent by making tiny changes that fall just outside the scope of a patent claim.¹⁸³ But finding infringement against an innocent innovator under the doctrine of equivalents can simply reward a patentee for writing fuzzy claims and relying on a default rule that requires the innovator to invest at its peril.¹⁸⁴

At this point, it is too soon to tell whether *Bilski* will survive Supreme Court review and what its full implications will be. On the one hand, much of what the Federal Circuit characterizes as the machine/transformation test in *Bilski* was developed by the Supreme Court in its *Benson* decision and prior decisions. The Court stated that the machine/transformation test was the “clue” to patentability.¹⁸⁵ *Bilski’s en banc* majority conducted a detailed analysis that located the machine/transformation test in Supreme Court precedent.¹⁸⁶

On the other hand are several technical problems and ambiguities. The machine/transformation test seems narrower and more rigid than the statutory language, which permits process patents and defines processes broadly, without reciting anything resembling the machine/transformation test.¹⁸⁷ Further, while *Bilski’s* concerns may be entirely legitimate, perhaps they are better expressed through other elements of the patenting process, such as the nonobvious requirement and the scope of prior art, the enablement requirement, or the doctrine of equivalents, all of which are intended to narrow the scope of patentable subject matter as broadly defined in 35 U.S.C. § 101. Stricter application of these doctrines could certainly yield more clarity and make the communication of patent boundaries more reliable.¹⁸⁸

At the same time, no part of the patent statute may be construed so as to make it inconsistent with either its constitutional authorization or its drafter’s intended purpose. If

¹⁸³ On this point, see Meurer & Nard, *supra* note 126.

¹⁸⁴ See John R. Allison & Mark A. Lemley, *The (Unnoticed) Demise of the Doctrine of Equivalents*, 59 STAN. L. REV. 955 (2007).

¹⁸⁵ See *Benson*, 409 U.S. at 70 (“Transformation and reduction of an article ‘to a different state or thing’ is the clue to the patentability of a process claim that does not include particular machines.”).

¹⁸⁶ *Bilski*, 545 F.3d at 953–59.

¹⁸⁷ See 35 U.S.C. § 101 (2006) (“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, subject to the conditions and requirements of this title.”); see also 35 U.S.C. § 100(b) (“The term ‘process’ means process, art, or method, and includes a new use of a known process, machine, manufacture, composition of matter, or material.”).

¹⁸⁸ See Michael Risch, *Everything is Patentable*, 75 TENN. L. REV. 591 (2008) (making this argument).

subject matter that does not meet the machine/transformation test falls outside of that authorization, it is no answer to say that the PTO examiner can subsequently fix things by applying a different technical doctrine. For example, nothing in section 101 of the Patent Act, which defines patentable subject matter, indicates that “abstract ideas” cannot be patented. The statute covers “any new and useful process,” making no exclusion for thought processes.¹⁸⁹ Nevertheless, the Supreme Court has consistently stated that abstract ideas cannot be patented.¹⁹⁰ With respect to laws of nature, also not mentioned in section 101, the Supreme Court has distinguished between the *discovery* of the laws themselves (unpatentable) and a particular application of those laws (patentable).¹⁹¹ This is so notwithstanding the fact that section 101 includes within its protection one who “*discovers* any new and useful process . . . or composition of matter.”¹⁹² Congress drafted section 101 broadly in order to extend patentable subject matter to the full range of its power under the IP Clause of the Constitution,¹⁹³ but that still gives the

¹⁸⁹ 35 U.S.C. § 101; *Metabolite Labs., Inc. v. Lab. Corp. of Am. Holdings*, 370 F.3d 1354 (Fed. Cir. 2004), *cert. dismissed*, 548 U.S. 124, 126–27 (2006) (Breyer, J., dissenting) (“The justification for the principle [that laws of nature are not patentable subject matter] does not lie in any claim that ‘laws of nature’ are obvious, or that their discovery is easy, or that they are not useful. To the contrary, research into such matters may be costly and time consuming; monetary incentives may matter; and the fruits of those incentives and that research may prove of great benefit to the human race. Rather, the reason for the exclusion is that sometimes *too much* patent protection can impede rather than ‘promote the Progress of Science and useful Arts,’ the constitutional objective of patent and copyright protection.”); *In re Comiskey*, 499 F.3d 1365, 1375 (Fed.Cir.2007) (“The Constitution explicitly limited patentability to ‘the national purpose of advancing the useful arts—the process today called technological innovation.’”). *cf.* *Parker v. Flook*, 437 U.S. 584, 596 (1978) (“[W]e must proceed cautiously when we are asked to extend patent rights into areas wholly unforeseen by Congress.”).

¹⁹⁰ *Diamond v. Diehr*, 450 U.S. 175, 185 (1981) (“Excluded from such patent protection are laws of nature, natural phenomena, and abstract ideas.”); *id.* at 196 (Stevens, J., dissenting) (“processes involving mental operations” not patentable subject matter); *Parker*, 437 U.S. at 598 (Stewart, J., dissenting) (“It is a commonplace that laws of nature, physical phenomena, and abstract ideas are not patentable subject matter.”). See also *Gottschalk v. Benson*, 409 U.S. 63, 67 (1972), which says the same thing but uses the term “abstract intellectual concepts.” *Cf.* *Rubber-Tip Pencil Co. v. Howard*, 87 U.S. (20 Wall.) 498, 507 (1874); *Le Roy v. Tatham*, 55 U.S. (14 How.) 156, 175 (1853) (“A principle, in the abstract, is a fundamental truth; an original cause; a motive; these cannot be patented . . .”).

¹⁹¹ *Diehr*, 450 U.S. at 187–88; *accord* *Funk Bros. Seed Co. v. Kalo Inoculant Co.*, 333 U.S. 127 (1948).

¹⁹² 35 U.S.C. § 101 (emphasis added). On the meaning of “discover,” or “discovery,” in the Constitution’s IP Clause and the Patent Act, and its relation to promotion of science and the useful arts, see Stephen McKenna, *Patentable Discovery?*, 33 SAN DIEGO L. REV. 1241 (1996), who notes that one important distinction between discoveries and inventions are that the discoveries already existed before they were discovered, even if they were not generally known. The author concludes that the distinction between “discoveries” (unpatentable) and “inventions” (patentable) is entirely judicial. In any event, both the framers of the Constitution and the drafters of section 101 probably regarded the terms “invention” and “discovery” as largely synonymous.

¹⁹³ *Diamond v. Chakrabarty*, 447 U.S. 303, 315 (1980) (“The subject-matter provisions of the patent law have been cast in broad terms to fulfill the constitutional and statutory goal of promoting ‘the Progress of Science and the Useful Arts’ with all that means for the social and economic benefits envisioned by Jefferson.”); *accord* *J. E. M. Ag Supply, Inc. v. Pioneer Hi-Bred Int’l, Inc.*, 534 U.S. 124, 131 (2001).

courts the job of determining what that power is.¹⁹⁴

Another critique is that *Bilski's* articulation of the machine/transformation test is incomplete, and too many details remain to be worked out. As a matter of fact, *Bilski's* analysis of the test and the precedents supporting it is lengthy and addresses most of the issues that are likely to arise. For example, once one uses a computer to calculate an algorithm, the data on the hard drive will not be exactly the same as before the calculation was made, but this clearly does not mean that the process “transformed” the computer in a qualifying way.¹⁹⁵ Nor does use of a lead pencil to perform the calculation “transform” the pencil by shortening the lead. *Bilski* made clear that these are not the kinds of transformations that create patent-worthy subject matter. “This transformation must be central to the purpose of the claimed process.”¹⁹⁶ Further, the court was clear that the use of a physical object must impose “meaningful limits” on the scope of a patent claim.¹⁹⁷ This limitation is essential in areas where abstract claims are capable of spilling over into activities that are completely unrelated to the subject matter of the patent.

There is also a bigger point here that should not be lost. Common law rules that answer every future question with finality are rare or non-existent. *Bilski*, like any important judicial decision, will require further clarification. The Federal Circuit acknowledged as much,¹⁹⁸ and so

¹⁹⁴ *Id.* at 115. See *In re Bilski*, 545 F.3d 943, 998 (Fed. Cir. 2008) (en banc) (Mayer, J., dissenting) (citing Pollack, *supra* note 153), *cert. granted*, 129 S. Ct. 2735 (2009) (No. 08-964); see also *Comiskey*, 554 F.3d at 977 (“The Constitution explicitly limited patentability to ‘the national purpose of advancing the useful arts—the process today called technological innovation.’”). Antitrust does very much the same thing that the Patent Act does. For example, the Sherman Act reaches to the full extent of Congressional power under the Commerce Clause. *United States v. South-Eastern Underwriters Ass’n*, 322 U.S. 533, 558 (1944). Nevertheless, the courts continue to decide whether jurisdiction under the commerce clause obtains in particular situations. See, e.g., *Summit Health, Ltd. v. Pinhas*, 500 U.S. 322 (1991) (divided Court finding jurisdiction); *United States v. Yellow Cab Co.*, 332 U.S. 218 (1947) (no jurisdiction). Other decisions are discussed in 3 AREEDA & HOVENKAMP, *supra* note 23, ¶266 (3d ed. 2006).

¹⁹⁵ See *Benson*, 409 U.S. at 64, 71, which concluded that simply adding a claim that an algorithm had to be implemented on a computer did not turn the algorithm into patentable subject matter; the claim was still on “the algorithm itself.” The court also noted that the calculation could be performed on computers “long in use” and thus no improvement in the computers themselves was being claimed. *Id.* at 67; see also *id.* at 71–72 (“It is conceded that one may not patent an idea. But in practical effect that would be the result if the formula for converting BCD numerals to pure binary numerals were patented in this case. The mathematical formula involved here has no substantial practical application except in connection with a digital computer, which means that if the judgment below is affirmed, the patent would wholly pre-empt the mathematical formula and in practical effect would be a patent on the algorithm itself.”).

¹⁹⁶ *Bilski*, 545 F.3d at 962.

¹⁹⁷ *Id.*

¹⁹⁸ See *id.* at 956 (recognizing “the possibility that this court may in the future refine or augment the test or how it is applied,” depending on “future developments in technology and the sciences”).

has the Supreme Court. In *Benson*, the Court stated a version of the machine/transformation test and used it to bar patentability of a mathematical formula whose only physical tie was that it was implemented on a computer.¹⁹⁹ At the same time, however, it stated that it did “not hold that no process patent could ever qualify if it did not meet the requirements of our prior precedents.”²⁰⁰ This language is sometimes read as the Supreme Court acknowledgement that tests other than the machine/transformation test that it described could be acceptable as well.²⁰¹ But that is not precisely what the Court said. Rather, it was saying that the Court’s “prior precedents” should not be read as determining the full set of definitions of patentable subject matter. If some technology came along that required expansion, distinction, or modification, the Court would take care of it. No judicial decision ever does any more.

Further, the machine/transformation analysis needs to be applied in a close case only when the other statutory elements of patentability are established.²⁰² The courts have sometimes spoken of section 101’s subject-matter provision as a “threshold” requirement that must be met for all inventions.²⁰³ This is certainly true, but nothing in the patent statute requires a patent examiner or a reviewing tribunal to look invariably to section 101 first, satisfy itself that the subject matter requirement has been met, and only then proceed to the other requirements for patentability. The most efficient course is to go to those issues that are most readily and unambiguously disposed of first. For example, if a patent clearly does not meet the section 103 requirement of nonobvious subject matter, the application can be rejected without further examination of any other factor of patentability. *Bilski* itself noted this.²⁰⁴ We do the same thing in antitrust law all the time. For example, the rule of reason in antitrust cases assesses substantial market power as a threshold requirement.²⁰⁵ However, if it is plain on the face of a complaint

¹⁹⁹ *Benson*, 409 U.S. 63.

²⁰⁰ *Id.* at 71; *see also* *Diamond v. Diehr*, 450 U.S. 175, 184 (1981) (quoting *Benson*’s “transformation and reduction to a different state or thing” language but not its qualifier about prior precedents).

²⁰¹ *See, e.g., Bilski*, 545 F.3d at 978 (Newman, J., dissenting).

²⁰² These are set forth in 35 U.S.C. § 102 (novelty, patented previously or elsewhere, described in a printed publication, on-sale bar), § 103 (non-obviousness, or inventive step), and § 112 (sufficient specification, enablement, and best mode).

²⁰³ *Diehr*, 450 U.S. at 188.

²⁰⁴ *In re Bilski*, 545 F.3d 943, 950 n.1 (Fed. Cir. 2008).

²⁰⁵ *E.g.,* *Theme Promotions, Inc. v. News Am. Mktg. FSI*, 546 F.3d 991, 1001 (9th Cir. 2008) (“A rule of reason analysis [for an antitrust claim] requires a threshold inquiry into the defendant’s market power.”); *Telecor Commc’ns, Inc. v. Sw. Bell Tel. Co.*, 305 F.3d 1124, 1148 (10th Cir. 2002) (“Market power is the preliminary threshold inquiry and is often dispositive of antitrust cases.”); *Hardy v. City Optical, Inc.*, 39 F.3d 765 (7th Cir. 1994) (“substantial market power is a threshold requirement of all rule of reason . . . cases”).

that the alleged conduct is not anticompetitive, then the court will dismiss the complaint without consideration of power.²⁰⁶

Bilski is clearly a step in the right direction. The machine/transformation test can go a long way toward addressing the problem of overly abstract patents, particularly of business methods. Excessive abstraction is a significant problem for competition policy and innovation because patents with excessively ambiguous boundaries create a property rights minefield. Perhaps *Bilski's* biggest and as yet unrealized contribution is as an aid to patent searchability. The description of a specific object or a specific way that an object will be transformed brings patent property much closer to traditional property concepts.

Before casting *Bilski* aside, it is useful to determine whether innovations can be imagined that should be patented but that do not qualify as patent-eligible subject matter under the machine/transformation test.²⁰⁷ If we can think of something, then we need to see if we can rewrite the test so as to include what we wish to include without bringing in too many other things that should be excluded. Insofar as the *Bilski* test creates a problem, it is excessive false negatives, or denials of patentability to patent-worthy subject matter. A better test would be one that can address the false negative problem without creating excessive false positives on the other side.

Further, looking prospectively rather than retrospectively is essential. It is too easy to criticize the test by looking backward and finding patents for inventions that we would deem patent-worthy but that do not satisfy the test. But those patent applications were drafted before the test was in place. The real importance of the machine/transformation test is the limitations that it imposes on those drafting patent applications in the future. The test serves as a limiting device on the drafter, requiring her to relate a patent's claims in some central way to tactile subject matter. Writing a patent application so as to satisfy the test is a discipline that serves to

²⁰⁶ See, e.g., *Cal. Computer Prods., Inc. v. IBM Corp.*, 613 F.2d 727, 744 (9th Cir.1979) (“assuming it was a monopolist,” defendant nevertheless did not violate monopolization law by engaging in precompetitive innovation).

²⁰⁷ One possibility is the fifth claim in the Morse telegraph patent, which referred to what has come to be known as Morse Code. See Brief of Regulatory Datacorp., Inc., et al., at 20, *Bilski v. Doll*, 129 S.Ct. 2735 (2009) (No. 08-964) (arguing that this claim was accepted even though it did not meet the machine/transformation test) That claim read:

‘Fifth. I claim, as my invention, the system of signs, consisting of dots and spaces, and of dots, spaces, and horizontal lines, for numerals, letters, words, or sentences, substantially as herein set forth and illustrated, for telegraphic purposes.

But this claim hardly referred to a simple alphabet with dots and dashes corresponding to each letter. Rather, the innovation was stated in the claim to be “for telegraphic purposes,” which signified that the claim was for a language that was to be implemented on a particular machine. The whole purpose of the system was to enable reception on an electric device that was capable of recognizing only short and longer bursts of electric current.

narrow the patent's reach and clarify its boundaries.

To be sure, we might hypothesize that some patent application that is unforeseeable today could not be written so as to satisfy *Bilski* and still seem eminently patent worthy. But a more open-ended test that creates numerous false positives and opens the door to heavy volumes of litigation is a poor way to account for that possibility. Antitrust law has acknowledged this fact more than once. For example, the law of predatory pricing requires strict cost-based tests, but not because we cannot conceive of plausible anticompetitive above-cost pricing strategies. Rather, it is because admitting such claims invites so many false positives that the welfare results would almost certainly be negative.²⁰⁸ In the case of *Bilski*, we do not even have a robust conception of a set of things that should be patentable but that cannot be written so as to satisfy the machine/transformation test.

3. *eBay: The Relation Between Remedies and Notice*

In its *eBay* decision, the Supreme Court rejected a Federal Circuit rule that gave patentees virtually blanket authority to obtain injunctions against infringement. Rather, the patentee must show the traditional set of equity requirements for an injunction -- namely, irreparable injury, inadequacy of damages as a remedy; and that the balance of hardships and the public interest favor an injunction.²⁰⁹

One might view *eBay* as creating compulsory licensing by another name—which might occur if courts simply permitted ongoing patent infringement with damages as the sole remedy. But *eBay* has its limitations as an alternative to compulsory licensing. First, it does not apply to *ex ante* requests for dealing but only to remedies once an infringement is found. This means that someone who knows about another's patent right in advance cannot insist on having it. Rather, the developer must go ahead and develop, risking that if infringement is found, a subsequent court will refuse to issue an injunction and award only damages. Further, knowing about a patent in advance may be a good reason for issuing the injunction—*eBay* has most of its bite in situations where developers did not know they were infringing and were caught by surprise.

So even if we thought that dominant firms had a duty to share newly developed technologies, *eBay* would not be a reliable way of implementing it.²¹⁰ As a result, it is not a

²⁰⁸ See HOVENKAMP, *supra* note 67, at 159–61.

²⁰⁹ *eBay Inc v. MercExchange, L.L.C.*, 547 U.S. 388, 391 (2006). The problem of *eBay* and patent trolls is discussed in Lee Anne Fennell, *Adjusting Alienability*, 122 HARV. L. REV. 1403 (2009).

²¹⁰ See Daniel A. Crane, *Intellectual Liability*, 88 TEX. L. REV. (forthcoming 2009) (manuscript available at <http://ssrn.com/abstract=1375031>) (arguing that the whole problem of compulsory licensing needs to be examined, and that mandatory licensing might serve as a policy offset for permitting other practices, such as bundling of IP

good substitute for an antitrust rule that compels dealing. *eBay* is better at dealing with various *ex post* problems that arise because patent boundaries are so poorly defined.²¹¹ While it is hardly a complete fix, *eBay* can operate to reduce the risk of innovating where patents held by others cannot effectively be searched in a way that is not prohibitively expensive.²¹² One good example is patents held by nonproducing entities, or patent “trolls,” who develop patents for complex technologies but do not practice them. Given the high costs and poor results of patent searches, a firm might discover only after the fact that it has made a costly, irreversible investment in technology covered by someone else’s patent.²¹³ One consequence of *eBay* is that courts have tended to limit nonpracticing patentees to damages.²¹⁴

As we argue below, dealing with patent trolls is not so much a problem of the right type of relief as the nature and timeliness of notice. Nonetheless, limiting the infringer’s liability to damages in cases where notice is poor serves to give innovators at least some confidence that they can move into a new area without subsequently facing prohibitive, gargantuan penalties that the right to an injunction can produce. A rule that placed a larger premium on notice would also give patentees an incentive to see to it that their patents are effectively communicated to potential infringers. For example, in *Rambus*²¹⁵ and *Qualcomm*²¹⁶ the patentee, who was participating in

rights, that would be less anticompetitive if there were a compulsory licensing duty).

²¹¹ See John Golden, Patent Remedies in Your Face, (Nov. 13, 2008), available at http://www.hbs.edu/units/tom/seminars/2007/docs/Introduction_to_Patent_Remedies_In_Your_Face.pdf.

²¹² See Mark A. Lemley, *Ignoring Patents*, 2008 MICH. ST. L. REV. 19 (cost of patent searches is so high, and results so unreliable, that firms often simply innovate now and worry about infringement litigation later). Nevertheless, the Federal Circuit has indicated that failure to search may result in a finding of willful infringement, with enhanced penalties up to treble damages. See *Clontech Labs., Inc. v. Invitrogen Corp.*, 406 F.3d 1347, 1357 n.6 (Fed. Cir. 2005).

²¹³ See Cotter, *supra* note 121 (manuscript at 32) (suggesting that denial of injunctive relief may be specially justified in situations involving nonpracticing patentees); Lemley & Shapiro, *supra* note 106, at 2036 (similar); Mark A. Lemley & Philip J. Weiser, *Should Property or Liability Rules Govern Information?*, 85 TEX. L. REV. 783, 798–99 (2007) (opposing injunctions in some situations). But the issue has been controversial. See also Einer Elhauge, *Do Patent Holdup and Royalty Stacking Lead to Systematically Excessive Royalties?*, 4 J. COMPETITION L. & ECON. 535, 537 (2008) (arguing in favor of preserving most injunctive relief); Damien Geradin, Anne Layne-Farrar & A. Jorge Padilla, *The Complements Problem Within Standard Setting: Assessing the Evidence on Royalty Stacking*, 14 B.U. J. SCI. & TECH. L. 144, 174–75 (2008) (similar); John M. Golden, “Patent Trolls” and Patent Remedies, 85 TEX. L. REV. 2111, 2148–49 (2007) (suggesting rebuttable presumption of injunctive relief).

²¹⁴ See Bernard H. Chao, *After eBay, Inc. v. MercExchange: The Changing Landscape for Patent Remedies*, 9 MINN. J. L. SCI. & TECH. 543 (2008); Michael W. Carroll, *Patent Injunctions and the Problem of Uniformity Cost*, 13 MICH. TELECOMM. & TECH. L. REV. 421 (2007); Vincenzo Denicolò, Damien Gerardin, Anne Layne-Farrar, & A. Jorge Padilla, *Revisiting Injunctive Relief: Interpreting eBay in High-Tech Industries with Non-Practicing Patent Holders* (Dec. 3, 2007) available at <http://ssrn.com/abstract=1019611>; see also Fennell, *supra* note 209.

²¹⁵ *Rambus Inc. v. FTC*, 522 F.3d 456 (D.C. Cir. 2008), *cert. denied*, 129 S. Ct.1318 (2009); see *supra* text

standard-setting deliberations, knew well in advance who the likely infringers would be and what would be the nature of their infringement. The socially preferable strategy is for them to communicate this to developers at a time when the developer can still make a choice. They chose to hold back, however, or even to misrepresent their holdings, in order to extract a larger award later. In such a case, the patentee should not be permitted to obtain more relief than damages measured by the value of substitute technologies measured *ex ante*—that is, prior to the time that a dedicated commitment was made.

On the other side is the question whether the infringer knew of the infringement at the time it developed its technology. If he did not know, then the case for an injunction is weak, particularly if the patentee was not producing or if patent enforcement is based on a continuation application that was pending when the infringer's technology was developed, and backdated to the date of original patent filing.²¹⁷ By contrast, the developer who knows in advance about another's patent but chooses to infringe it anyway is in a much different position. In that case, the award of an injunction is more appropriate because the developer could have negotiated a license in advance, when the full range of alternatives was still available. Further, whether or not the patentee is practicing the patent does not matter all that much.

In sum, the remedy system should be designed so as to reward patentees who give adequate and timely notice of their claims, and to penalize those who do not. It should also reward innocent infringers and penalize those who proceeded to develop even though they knew or suspected they would be infringers.

None of this is intended as a brief for either side of the question whether property or liability rules should protect IP owners.²¹⁸ Those who want to preserve broad property rights generally emphasize the systematic undercompensation they believe results from judicial determinations of damages, and the resulting harm to the incentive to innovate.²¹⁹ Otherwise they are concerned about the market integrity of prices determined by judges rather than buyers and sellers.²²⁰ Those who favor relaxing injunction entitlements are concerned mainly with

accompanying notes 120-121.

²¹⁶ *Qualcomm Inc. v. Broadcom Corp.*, 548 F.3d 1004 (Fed. Cir. 2008), *cert denied*, 129 S. Ct. 2182 (2009).

²¹⁷ On the problem of patent continuations, see *supra* text accompanying notes 125-130.

²¹⁸ See Crane, *supra* note 210.

²¹⁹ E.g., Richard A. Epstein, *The Property Rights Movement and Intellectual Property*, 30 REGULATION 58, 62 (2008) (“[S]ystematic under-compensation during the limited life of a patent is likely to reduce the level of innovation while increasing the administrative costs of running the entire system.”); see also Elhauge, *supra* note 213.

²²⁰ Henry E. Smith, *Property and Property Rules*, 79 N.Y.U. L. Rev. 1719 (2004).

holdup and excessive royalties.²²¹

The concern about chronic undercompensation has merit, but mainly in a system in which property rights are well defined and notice to outsiders is adequate and timely. And holdup is surely of concern when infringers are taken by surprise in circumstances where an objectively reasonable person would not have known about the right. But the availability of remedies can be metered in a way that minimizes both concerns. When IP rights holders have taken all reasonable steps to communicate their rights to others, including specific and timely communications to those it knows are likely to infringe, then the case for equity relief is much stronger than when they keep silent while knowing of likely infringement, or take other steps to withhold information about their rights.²²² On the other side, an infringer who had good reason to know it was infringing at the time it made its investment should not be heard later to complain that it is being held up by a rapacious troll seeking to appropriate its investment. But an infringer who acted reasonably and had no reason to suspect infringement until it is too late makes a good case for a damages remedy. As the infringement itself moves further away from anything that was foreseeable to the innovator at the time of its innovation, perhaps because the claims were overly abstract or were the unpublished contents of a continuation application, the case for an injunction grows much weaker.

To be sure, not every bargaining problem goes away when people have timely knowledge of property rights. For example, bilateral monopoly problems exist even when property rights are well defined.²²³ Such problems are likely to arise when a standard-setting organization dominates its market and the patentee has a clearly superior technology to offer. The two sides are then thrust into a situation where they can deal only with each other, or at least where offers from everyone else are clearly inferior.

Ultimately, neither antitrust policy nor IP policy may be able to solve bilateral monopoly problems. But bilateral monopoly is ubiquitous and hardly limited to IP rights.²²⁴ Of course,

²²¹ Peter Lee, *The Evolution of Intellectual Infrastructure*, 83 WASH. L. REV. 39 (2008). And see the following exchange: Lemley & Shapiro, *supra* note 110; Golden, *supra* note 213; Mark A. Lemley & Carl Shapiro, *Reply: Patent Holdup and Royalty Stacking*, 85 TEX. L. REV. 2163 (2007).

²²² The penalty of judicially determined royalties rather than an injunction for the patentee who fails to provide effective, timely notice of its rights strikes some as severe. See, e.g., Elhauge, *supra* note 213. But it should be kept in mind that the penalty to the real property owner who fails to provide adequate notice is often complete loss of title. And where search costs are higher, the property owner should have an even stronger obligation to provide notice.

²²³ Roger D. Blair et al., *A Pedagogical Treatment of Bilateral Monopoly*, 55 S. ECON. J. 831 (1989).

²²⁴ See Fennell, *supra* note 209 (discussing problems of bilateral monopoly and proposing revision of alienability rules to address them); Herbert Hovenkamp, *Arthur Cecil Pigou and the Coase Theorem*, ___ ARIZ. L. REV. ___ (2009) (in press) (on bilateral monopoly in Coasian markets).

buyers' cartels are unlawful when they are found, and antitrust's rule of reason is designed for such situations. A standard-setting organization that is nothing more than a front for a cartel ought to be condemned under the antitrust laws. However, one that is engaged in true "joint purchasing" of technology that will be used to facilitate a valuable common standard will be approved even if it has power in the buying market.²²⁵ Buyer power problems are likely to arise when a group of firms that dominate a market set a mandatory standard, as is likely to be the case of things such as network compatible communications devices.²²⁶ And in such cases, we may have to countenance a legitimate buying organization that is able to suppress the price of the technology that it purchases. But that is hardly an argument for the patentee's concealment of its rights until the buyers have made their investment.

"Judicial" patent reform is underway, and the problems of overbreadth, notice and priority are significantly less imposing today than they were a decade ago. At the same time, *KSR*, *Bilski*, and *eBay* are all relatively recent decisions. At this writing it is difficult to predict what their full impact will be.

C. Legislative and Regulatory Sources of Patent Reform

The impetus for patent reform is also coming from other sources. Recently, the PTO itself initiated several reforms. One of the most important ones, which a panel of the Federal Circuit subsequently rejected, was severe limitations on patent continuations.²²⁷ Under the proposed regulations, a patent applicant would have been entitled to two continuation applications subsequent to the original application. If a third application was filed, the applicant would have to show that its claims could not have been stated in the previous applications. Otherwise, the third application would be treated as a new application and given its own filing date for priority purposes. The Federal Circuit struck this rule down as a violation of the patent statute's requirement that the validity date of a patent relate back to its initial filing date.²²⁸

The 2009 Patent Reform Act, currently pending before Congress,²²⁹ moves the United

²²⁵See Herbert Hovenkamp, *Standards Ownership and Competition Policy*, 48 B.C. L. REV. 87 (2007).

²²⁶ *Golden Bridge Tech., Inc. v. Motorola, Inc.*, 547 F.3d 266 (5th Cir. 2008) (standard-setting organization for cellular phones acted reasonably in setting a standard that excluded the plaintiff's technology), *cert. denied*, 129 S. Ct. 2055 (2009).

²²⁷See discussion *supra*, text at notes 120-132.

²²⁸ *Tafas v. Doll*, 559 F.3d 1345 (Fed. Cir. 2009). The Federal Circuit has granted a rehearing *en banc*. No. 2008-1352, 2009 WL 1916498 (Fed. Cir. July 6, 2009).

²²⁹ Patent Reform Act of 2009, H.R. 1260, 111th Cong. (2009).

States from a “first to invent” to a “first to file” system for determining priority, places limits on patent damages to the “specific contribution over prior art” that the patented invention made to the infringer’s technology, and provides for expanded reexamination processes with increased opposition allowed by third parties. However, the bill does not address most of the problems described in this Section. Further, Congress has been beset by lobbying over patent legislation, and previous patent reform bills have failed.²³⁰ As a result, the most significant patent reforms are likely to come not from statutory amendments or even from PTO rulemaking, but rather from the courts themselves.

IV. Copyright Law’s Errand

The situation for copyright is bleaker than for patents. While core copyright protection is arguably necessary to allow authors and other copyright holders prevent harmful piracy of their works, the 1976 Copyright Act confers excessively broad copyrights that prohibit most copying of copyrighted material, even when the copying does no harm to the copyright holders’ incentives to create or distribute their works. Indeed, the Act shows the signs of special interest capture even more than the Patent Act does: it benefits well-organized interest groups while imposing costs on the general public; it contains numerous special provisions favoring this or that interest group; the legislative history shows extensive interest group involvement; and the scope of statutory rights is indefensible on public interest grounds.²³¹ Yet, unlike in patent law, there is no reform on the horizon for copyright law.

As Jessica Litman has written, “Unlike the porous 1909 [Copyright] Act, the 1976 Act is a detailed comprehensive code, chock-full of specific heavily negotiated compromises.”²³² The current statute is generally characterized by broad rights for copyright holders and narrow, specific exemptions for users. For instance, the copyright holder’s right of public performance generally excludes others from “performing” a work in public.²³³ As the right is defined, it is broad enough to cover singing a song to oneself while walking through a public shopping mall, although one of the specifically drawn exemptions to this right might ultimately defeat a finding of infringement.²³⁴

²³⁰ See Patent Reform Act of 2008, S. 3600, 110th Cong. (2008); Patent Reform Act of 2007, S. 1145, 110th Cong. (2007); H.R. 1908, 110th Cong. (2007); Patent Reform Act of 2006, S. 3818, 109th Cong. (2006); Patent Reform Act of 2005, H.R. 2795, 109th Cong. (2005).

²³¹ Bohannan, *supra* note 56, at 568 (showing that the Copyright Act bears all the indicia of special-interest legislation).

²³² Jessica D. Litman, *Copyright, Compromise, and Legislative History*, 72 CORNELL L. REV. 857, 859 (1987).

²³³ See 17 U.S.C. § 106(4) (2006).

²³⁴ See *id.* § 110(4) (exempting the noncommercial performance of a nondramatic literary or musical work). The other statutory exemptions to the public performance right are narrowly drawn to exempt activities such as “face-to-face” teaching in a non-profit educational institution or for performing a non-dramatic literary or musical work of a

Moreover, the statute grants to copyright holders the exclusive rights not only to copy but also to “modify, transform, or adapt” the copyrighted work in any way.²³⁵ Put together, these broad rights prohibit most copying of copyrighted material unless some exception applies.²³⁶ What is more, these rights endure for several generations: for the life of the author plus seventy years in the case of a work by a natural author, or for ninety-five years in the case of a work for hire.²³⁷

The scope and duration of these rights far exceed what is necessary to encourage the creation and dissemination of copyrighted works. Although the goal of the IP Clause of the Constitution is to “promote the Progress of Science” by allowing Congress to grant copyrights only “for limited Times,”²³⁸ the current copyright term is the economic equivalent of indefinite ownership.²³⁹ The scope of copyrights is similarly problematic. A copyright holder can enjoin or demand royalties for virtually all copying, regardless of whether that copying was of a kind likely to harm the copyright holder’s incentives to innovate.

To be sure, the fair use doctrine considers “harm to the market for the copyrighted work” as one factor in determining whether a use is fair. There are, however, numerous problems with its approach. First, harm is not a strict requirement but is merely one factor to be considered in the analysis.²⁴⁰ Second, because fair use is treated as an affirmative defense, the burden of proof is on

religious nature in the course of worship. *See id.* § 110.

²³⁵ *See id.* § 106(2).

²³⁶ *See* LANDES & POSNER, *ECONOMIC STRUCTURE*, *supra* note 74, at 110 (arguing that “since it is uncertain whether *any* copyright protection, let alone the amount conferred by current law, is necessary to enable authors and publishers to recover the fixed costs that must be incurred to generate the socially optimal output of expressive works, it would be speculative to conclude that without control over derivative works authors and publishers would be unable to cover those costs.”).

²³⁷ *See* 17 U.S.C. § 302(a), (c) (2006).

²³⁸ U.S. CONST. art. I, § 8, cl. 8; *see also* *Eldred v. Ashcroft*, 537 U.S. 186, 260 (2003) (Breyer, J., dissenting) (observing that the Patent and Copyright Clause “assumes that it is the *disappearance* of the monopoly grant, not its *perpetuation*, that will, on balance, promote the dissemination of works already in existence”).

²³⁹ *See Eldred*, 537 U.S. at 255–56 (Breyer, J., dissenting) (“The present extension will produce a copyright period of protection that, even under conservative assumptions, is worth more than 99.8% of protection *in perpetuity* (more than 99.99% for a songwriter like Irving Berlin and a song like Alexander’s Ragtime Band). The lack of a practically meaningful distinction from an author’s *ex ante* perspective between (a) the statute’s extended terms and (b) an infinite term makes this latest extension difficult to square with the Constitution’s insistence on ‘limited Times.’” (citation omitted)).

²⁴⁰ *See* 17 U.S.C. §107 (listing four factors courts must consider in fair use analysis); *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 590-91 (1994) (discussing fair use factors and saying that no single factor is dispositive).

the defendant to show the *absence* of harm.²⁴¹ Third, harm is so poorly defined in the case law that the concept has become circular. Because the statute grants copyright holders virtually complete control over copying of their copyrighted works, a copyright holder can always argue that the defendant's copying caused her harm because the defendant could have paid her a license fee for the very copying that the defendant claims is fair.²⁴² The effects of this circularity are far-reaching. It causes risk-averse users to obtain licenses even when they are unnecessary, and then the existence of those licenses tends to reify the notion that the right to control such uses exists.²⁴³ Moreover, courts hold that the copyright holder is entitled to licensing fees even for uses that *increase* demand for the original copyrighted work – that is, for uses that are effectively complements to rather than substitutes for the copyrighted work.²⁴⁴

The fundamental problem here is that the harm element in fair use is not tied to copyright's purpose of encouraging innovation. While a court might consider whether the defendant should compensate a particular copyright holder for a particular use, the court does not inquire into whether the copyright holder would have relied on compensation for that use in deciding whether to create or distribute the copyrighted work.

Courts have found infringement or likely infringement in numerous cases where harm to the copyright holder's incentives was nonexistent, speculative, or trivial. For instance, a district court preliminarily enjoined a book entitled *The Wind Done Gone* by Alice Randall, concluding that success on the fair use defense was unlikely.²⁴⁵ Randall's book borrowed characters and plot lines, among other things, from *Gone With the Wind* in order to produce a new story that critiqued the

²⁴¹ See *Campbell*, 510 U.S. at 590–91 (1994) (defendant bears burden of proof on fair use).

²⁴² See *Princeton Univ. Press v. Mich. Document Servs.*, 99 F.3d 1386, 1387 (6th Cir. 1996) (discussing circularity problem); *Am. Geophysical Union v. Texaco*, 60 F.3d 913, 929 (2d Cir. 1994) (same); see also Bohannon, *supra* note 85, at 978–79; Mark A. Lemley, *Should a Licensing Market Require Licensing?*, 70 LAW & CONTEMP. PROBS. 185, 190–91 & nn. 31–40 (2007); Matthew Africa, Comment, *The Misuse of Licensing Evidence in Fair Use Analysis: New Technologies, New Markets, and the Courts*, 88 CAL. L. REV. 1145, 1174–75 (2000).

²⁴³ See James Gibson, *Risk Aversion and Rights Accretion in Intellectual Property Law*, 116 YALE L.J. 882, 884 (2007) (“the practice of licensing within gray areas eventually makes those areas less gray, as the licensing itself becomes the proof that the entitlement covers the use”).

²⁴⁴ See Bohannon, *supra* note 56, at 596–97 (discussing examples of derivative works that increase sales of the original works on which they are based); see also *Ty, Inc. v. Publ'ns Int'l Ltd.*, 292 F.3d 512, 517–19 (7th Cir. 2002) (distinguishing between uses of copyrighted material that are economic substitutes for the original work and uses that are economic complements).

²⁴⁵ See *Suntrust Bank v. Houghton Mifflin Co.*, 136 F.Supp.2d 1357, 1386 (N.D. Ga. 2001), *vacated*, 252 F.3d 1165 (11th Cir. 2001), and *rev'd*, 268 F.3d 1257 (11th Cir. 2001).

classic story by re-telling portions of it from the point of view of a young slave girl.²⁴⁶ Despite acknowledging the critical and parodic character of *The Wind Done Gone*, and with little or no supporting evidence involving likely market effects, the district court held that the new work was likely to substitute for authorized sequels of *Gone With the Wind*.²⁴⁷

The Eleventh Circuit reversed the injunction.²⁴⁸ Its review of the record disclosed that “Suntrust focuses on the value of [*Gone With the Wind*] and its derivatives, but fails to address and offers little evidence or argument to demonstrate that [*The Wind Done Gone*] would supplant demand for Suntrust’s licensed derivatives.”²⁴⁹ Indeed, since the preliminary injunction was reversed and the new book went on sale, anecdotal market evidence suggests that the book actually *enhanced* sales of the original and its derivatives. Amazon.com reports that customers who have purchased the new book also have purchased the original *Gone With the Wind* novel, its authorized sequel *Scarlett*, and the authorized motion picture.²⁵⁰

More recently, another district court entered a permanent injunction against the sale and distribution of *Lexicon*, an encyclopedic reference guide to the *Harry Potter* series.²⁵¹ The court first held that plaintiffs J.K. Rowling and Warner Brothers did not have the right to control the market for reference guides to their works, and therefore that the *Lexicon* did not cause any cognizable harm to that market.²⁵² The court also concluded that the reference guide was not likely to supplant sales of the original seven *Harry Potter* novels.²⁵³ The court opined, however, that the

²⁴⁶ See *id.* at 1370-71 (noting that *The Wind Done Gone* “is principally and purposefully a critical statement that seeks to rebut and destroy the perspective, judgments, and mythology of [*Gone With the Wind*],” and that “[w]here Randall refers directly to Mitchell’s plot and characters, she does so in service of her general attack” on the book).

²⁴⁷ See *id.* at 1383.

²⁴⁸ See *Suntrust Bank v. Houghton Mifflin Co.*, 268 F.3d 1257 (11th Cir. 2001).

²⁴⁹ See *id.* at 1274-75 (adding that “the evidence proffered in support of the fair use defense specifically and correctly focused on market substitution and demonstrates why Randall’s book is unlikely to displace sales of [*Gone With the Wind*]”).

²⁵⁰ See http://www.amazon.com/Wind-Done-Gone-Novel/dp/0618219064/ref=sr_1_1?ie=UTF8&s=books&qid=1250356777&sr=1-1 (last visited August 15, 2009).

²⁵¹ See *Warner Bros. Entm’t, Inc. v. RDR Books*, 575 F.Supp.2d 513 (S.D.N.Y. 2008).

²⁵² See *id.* at 549-50 (“Notwithstanding Rowling’s public statements of her intention to publish her own encyclopedia, the market for reference guides to the *Harry Potter* works is not exclusively hers to exploit or license, no matter the commercial success attributable to the popularity of the original works.”).

²⁵³ See *id.* at 550 (“Children may be an elusive market for book publishers, but it is hard to believe that a child, having read the *Lexicon*, would lose interest in reading (and thus his or her parents’ interest in purchasing) the *Harry*

Lexicon might potentially harm sales of Rowling's two companion books *Quidditch Through the Ages* and *Fantastic Beasts & Where to Find Them*.²⁵⁴ Yet, there was no evidence of such harm, and it seems inappropriate to presume that an encyclopedic reference guide would supplant demand for Rowling's own fictional books, which look and read very differently. For instance, the *Fantastic Beasts* book purports to be Harry's own course book for his Magical Creatures class at Hogwarts and contains Harry's own amusing hand-written notes throughout. What is more, the court speculated that "[a]lthough there is no supporting testimony," the *Lexicon* reference guide could possibly have an effect on future works, such as musical productions or "print publications of [the] songs and poems" that appear in the *Harry Potter* books.²⁵⁵ But the plaintiffs submitted no evidence indicating that they planned to enter those markets, or that, if they did, the *Lexicon* would have any negative effect on those markets whatsoever. Indeed, it seems just as likely that the *Lexicon* reference guide would enhance sales of Rowling's works in current and future markets.

In these cases and others like them, courts prohibit or discourage many uses of copyrighted works absent proof that the defendant's use is likely to harm the plaintiff's incentives to produce copyrighted works.²⁵⁶ This is a particular problem in cases involving transformative works, where others innovate by improving on existing works. At one time, such transformative works were themselves considered valuable "works of authorship" and, as such, were held to be noninfringing.²⁵⁷ As Landes and Posner have observed, these transformative uses often require

Potter series. . . . The *Lexicon* is thus unlikely to serve as a market substitute for the *Harry Potter* series and cause market harm.").

²⁵⁴ The court reasoned that "[u]nless they sought to enjoy the companion books for their entertainment value alone, consumers who purchased the *Lexicon* would have scant incentive to purchase either of Rowling's companion books, as the information contained in these short works has been incorporated into the *Lexicon* almost wholesale." See *id.* Of course, the same could be said of *Lexicon*'s encyclopedic summaries of the seven original *Harry Potter* novels, but the court concluded that there was no likely harm to the market for those original novels. Clearly, people do read Rowling's books for the entertainment value, not merely to obtain the information therein. In any event, the point here is simply that courts should not presume harm in such cases but should require evidence showing that the defendant's copying will actually reduce the copyright holder's incentives.

²⁵⁵ See *id.* at 551.

²⁵⁶ See, e.g., *Princeton Univ. Press v. Mich. Doc. Serv., Inc.*, 99 F.3d 1381, 1386-87 (6th Cir. 1996) (copyshop's copying of college coursepacks not fair use although professors testified that they would not have purchased copyrighted works in the absence of fair use and no other evidence indicated that copyright holders were worse off as a result of the copying); *Ty, Inc. v. West Highland Publ'g Inc.*, 1998 WL 698922, at *16 (N.D. Ill. 1998) (granting preliminary injunction against defendant's publication of reference guide to Ty's Beanie Babies stuffed animals based on speculation that the guide might have a negative effect on their "marketing image" at some point in the future).

²⁵⁷ See BENJAMIN KAPLAN, AN UNHURRIED VIEW OF COPYRIGHT 10 (1966) (observing that as of the early 19th century, the rule seemed to be that "if the accused book was a work of authorship, it could not at the same time infringe").

substantial investment on the part of the improver, and therefore do not create the same free-riding problems that simple verbatim copying does.²⁵⁸ Moreover, many uses of copyrighted material are not only innovative, but also constitute protected speech.²⁵⁹ Thus, when copyright law prohibits the use of copyrighted expression absent proof of harm to copyright incentives, it creates a glaring exception to the First Amendment rule that government may not prohibit speech absent a showing that the speech causes harm to a significant governmental interest.²⁶⁰

Evidence of copyright reform is very hard to find. To the contrary, in recent years, copyright law has been operating as a one-way ratchet favoring holders over consumers. On the legislative front, in 1998 Congress enacted the Copyright Term Extension Act (CTEA), which extended the term of future and existing copyrights by 20 years. The Supreme Court later upheld retroactive application of this term extension to existing copyrighted works, despite the fact that retroactive term extensions to existing works did little or nothing to promote innovation, while they significantly burdened future use and innovation by others.²⁶¹ Recent copyright legislation has expanded copyrights even further. For instance, in the past few years, Congress has resurrected copyrights for many works that previously had fallen into the public domain.²⁶²

Unfortunately, unlike in patent and antitrust law, there is very little evidence of copyright reform in the courts. The Supreme Court has not been consistent in making copyright protection commensurate with copyright's purpose. The Court's 1984 decision in *Sony Corp. of Am. v. Universal City Studios, Inc.*²⁶³ recognized the centrality of harm and the need to prove it. Writing

²⁵⁸ See LANDES & POSNER, *ECONOMIC STRUCTURE*, *supra* note 74, at 109-115. See also *Mills Music, Inc. v. Snyder*, 469 U.S. 153, 174 (1985) (noting that many firms make large investments in derivative works).

²⁵⁹ See Christina Bohannon, *Copyright Infringement and Harmless Speech* (Univ. of Iowa Legal Studies Research Paper, Paper No. 09-15, 2009), available at <http://ssrn.com/abstract=1367624> (arguing that because copying often constitutes speech, the First Amendment requires that applications of copyright law be justified by a showing that the copying harms the copyright holder's incentives to create or distribute copyrighted works); Rebecca Tushnet, *Copy This Essay: How Fair Use Doctrine Harms Free Speech and How Copying Serves It*, 114 *YALE L.J.* 535 (2004) (arguing that both transformative and non-transformative copying has speech value).

²⁶⁰ See Bohannon, *Copyright Infringement and Harmless Speech*, *supra* note 260.

²⁶¹ See *Eldred v. Ashcroft*, 537 U.S. 186, 249-251 (2003) (Breyer, j., dissenting).

²⁶² See Uruguay Round Agreement Acts (URAA) § 514, Pub. L. No. 103-465, 108 Stat. 4809, 4976-80 (1994) (current version at 17 U.S.C. §§ 104(A), 109 (2006)) (restoring U.S. copyright protection to certain foreign works that were previously in the public domain in the U.S.). This legislation has produced mixed results in the courts. *Compare* *Golan v. Gonzales*, 501 F.3d 1179 (10th Cir. 2007) (holding provision unconstitutional), *with* *Luck's Music Library, Inc. v. Gonzales*, 407 F.3d 1262 (D.C. Cir. 2005) (holding provision constitutional).

²⁶³ 464 U.S. 417 (1984).

for the majority, Justice Stevens explained:

The purpose of copyright is to create incentives for creative effort. . . . [A] use that has no demonstrable effect upon the potential market for, or the value of, the copyrighted work, need not be prohibited in order to protect the author's incentive to create. . . . Actual present harm need not be shown; such a requirement would leave the copyright holder with no defense against predictable damage. . . . What is necessary is a showing by a preponderance of the evidence that some meaningful likelihood of future harm exists.²⁶⁴

Sony went on to hold that home copying of copyrighted television programs for purposes of “time-shifting,” or watching the programs at a later time, constituted fair use because the evidence showed that “[h]arm from time-shifting is speculative and, at best, minimal.”²⁶⁵

Yet, the Court's more recent decision in *Campbell v. Acuff-Rose Music*²⁶⁶ seems to retreat from *Sony*. In *Campbell*, the Court held that the defendant 2 Live Crew's rap version of Roy Orbison's song *Oh, Pretty Woman* reasonably could be perceived as a parody of the original song.²⁶⁷ The Court sensibly concluded that copyright holders do not control the market for parodies and other criticism of their works, and therefore that the unauthorized parody did not cause cognizable harm to that market under the Copyright Act.²⁶⁸ Although the Court did inquire into harm, it seemed less emphatic about the need to limit the scope of copyrights to protect only against the kind of harmful copying that is likely to reduce incentives to create and distribute copyrighted works. First, the Court stated that copyright holders do get to control satirical uses of their works, despite the difficulty of distinguishing between parody and satire and the seemingly low probability that copyright holders would rely on the market for satirical uses of their works in deciding whether to create or distribute their works. Second, the Court suggested that the defendant might lose on its fair use defense if its parody caused any harm to the remote market for non-parody rap versions of the original song, and it remanded for further findings on this issue.²⁶⁹ Third, although there was no

²⁶⁴ *Sony*, 464 U.S. at 450-51 (internal citations omitted).

²⁶⁵ *Id.* at 454 (quoting the district court's decision at *Universal Studios, Inc. v. Sony Corp. of Am.*, 480 F. Supp. 429, 469 (C.D. Cal. 1979)). *See also* Bohannon, note 85, at 991-96 (showing *Sony*'s harm-based approach to fair use).

²⁶⁶ 510 U.S. 569 (1994).

²⁶⁷ *See id.* at 583.

²⁶⁸ *See id.* at 592-93.

²⁶⁹ *See id.* at 593-94.

reason to presume on the facts that any harm had occurred or was likely to occur, the Court held that the defendant bore the burden to show the *absence* of harm as part of its fair use defense.²⁷⁰ The Supreme Court's ambivalence in these cases suggests that it has not felt the same urgency as it has in patent law to bring copyright law into harmony with its constitutional justification.

V. Toward a Concept of "IP Injury"

Recent reforms, particularly in patent law, are a promising step toward encouraging innovation and competition. Much remains to be done. Landes and Posner have argued that the common law can provide useful baselines for measuring the proper scope of IP entitlements because the common law is more insulated from special interests. As a result, judges are more likely to keep their eye on the ball.²⁷¹ We agree, and believe that IP can profitably borrow from the "common law" of antitrust as well. Mainly, we argue for a conception of IP harm that would borrow from the "antitrust injury" doctrine and antitrust's strict, common-law-driven requirements that the right kind of harm and damages be proven. In IP law, provable harm must relate to the incentive to innovate, just as in antitrust provable harm relates to the incentive to compete. By querying first how the plaintiff was injured, a court might avoid much more difficult decisions concerning the appropriate scope of patent or copyright protection. By insisting on proof of the right kind of harm, the courts can often avoid much more intractable questions about the scope of coverage.

The antitrust laws are enforced by a mixture of public and private actions. Both the Antitrust Division of the Department of Justice and the Federal Trade Commission enforce them, but the great majority of actions are brought by private parties.²⁷² With rare exceptions the IP laws are enforced almost exclusively by private parties. Much of the revolution that has gone on in antitrust over the past thirty years has pertained to private actions and the types of harm needed to support them. Indeed, the antitrust revolution was driven in large part by the idea that private parties use the antitrust laws for anticompetitive ends, mainly to protect themselves from intense competition.²⁷³ As we have shown above, the same is true of IP law and the incentive to innovate.

The doctrine that has become the poster child for antitrust reform is the "antitrust injury"

²⁷⁰ See *id.* at 590.

²⁷¹ LANDES & POSNER, *supra* note 74, at 10, 87, 205–06.

²⁷² In most years, around ninety percent of antitrust complaints are filed by private parties. See BUREAU OF JUSTICE STATISTICS, U.S. DEP'T OF JUSTICE, SOURCEBOOK OF CRIMINAL JUSTICE STATISTICS ONLINE tbl. 5.41.2008, <http://www.albany.edu/sourcebook/pdf/t5412008.pdf> (last visited July 30, 2009).

²⁷³ *E.g.*, Baumol & Ordover, *supra* note 13; Easterbrook, *supra* note 13; William H. Page, *Optimal Antitrust Penalties and Competitors' Injury*, 88 MICH. L. REV. 2151 (1990); Snyder & Kauper, *supra* note 13.

rule developed by the Supreme Court in its 1977 *Brunswick* decision.²⁷⁴ The *Brunswick* case is illustrative, both for how it relates its harm requirement to the goals of the antitrust laws, and how it imposes a limiting interpretation on a private enforcement provision that seems clear and expansive on its face. The plaintiff was a struggling, independent bowling alley in Pueblo, Colorado. It faced one competitor, which was financially even worse off than the plaintiff. The competitor was also deeply in debt to its principal hardware supplier, Brunswick, a very large firm that produced an array of sporting goods including equipment for bowling alleys. The rival bowling alley settled its debts under an arrangement in which Brunswick acquired the alley, gave it an infusion of cash, and began operating it as a subsidiary. The plaintiff then filed suit, claiming that Brunswick's acquisition of the struggling alley was an unlawful merger. As a result, the plaintiff claimed, instead of being in a market where it faced a listless competitor on the verge of bankruptcy, it had to face a newly energized rival with substantial resources from above.

The facts immediately invite a double take. The merger may or may not have been unlawful, depending on the market shares of the merging firms and the height of entry barriers.²⁷⁵ Further, the plaintiff was clearly "injured" by the merger. Being forced to compete with a rejuvenated, Brunswick-owned facility would be much tougher than competing with a debt-ridden independent. Thus, under a plain reading of the expansive private-action provision of the antitrust laws, which requires only injury and causation,²⁷⁶ all of the statutory language appeared to favor the plaintiff's position, provided that it could show that the merger was unlawful.

But the Supreme Court did not focus on whether the merger was technically unlawful. Indeed, it did not even reach the issue. The only thing that mattered was that the plaintiff was complaining about *more* competition rather than less in the Pueblo bowling market. The nature of his complaint was that he would have been a dominant firm with a failing competitor, and that the merger forced him to face a much more aggressive rival. Clearly, recognizing such a complaint would be inimical to the goals of the antitrust laws.

The Supreme Court responded with a rule that requires a private antitrust plaintiff to show not just any injury, but *antitrust* injury—that is, injury that results from decreased competition. The rule was fashioned by one of the Court's great, pro-antitrust liberals and a survivor of the Warren Era, Justice Thurgood Marshall. Further, the antitrust-injury rule that he formulated seems quite inconsistent with the statutory language, which provides that: "[A]ny person who shall be injured in his business or property by reason of anything forbidden in the antitrust laws may sue therefor

²⁷⁴ *Brunswick Corp. v. Pueblo Bowl-O-Mat, Inc.*, 429 U.S. 477 (1977).

²⁷⁵ On these substantive requirements of merger law, see 4 & 4A AREEDA & HOVENKAMP, *supra* note 5, chs. 9–10 (3d ed. 2009).

²⁷⁶ 15 U.S.C. § 15 (2006).

. . . and shall recover threefold the damages by him sustained”²⁷⁷ That language seems to be both mandatory and comprehensive: all it requires is an injury to business or property that is caused by an antitrust violation. Nevertheless, Justice Marshall wrote for a unanimous Court:

Every merger of two existing entities into one, whether lawful or unlawful, has the potential for producing economic readjustments that adversely affect some persons. But Congress has not condemned mergers on that account; it has condemned them *only when they may produce anticompetitive effects*. Yet under the Court of Appeals’ holding, once a merger is found to violate §7, all dislocations caused by the merger are actionable, regardless of whether those dislocations have anything to do with the reason the merger was condemned. . . .

... At base, respondents complain that by acquiring the failing centers petitioner preserved competition The damages respondents obtained are designed to provide them with the profits they would have realized had competition been reduced.

. . . .
. . . [T]o recover treble damages on account of §7 violations, [plaintiffs] must prove more than injury causally linked to an illegal presence in the market. Plaintiffs must prove *antitrust* injury, which is to say injury of the type the antitrust laws were intended to prevent and that flows from that which makes defendants’ acts unlawful. *The injury should reflect the anticompetitive effect either of the violation or of anticompetitive acts made possible by the violation*. It should, in short, be “the type of loss that the claimed violations . . . would be likely to cause.”²⁷⁸

The Court’s decision in *Brunswick* revolutionized private antitrust enforcement, and lack of “antitrust injury” has led to the dismissal of hundreds of private antitrust cases.²⁷⁹ Indeed, antitrust injury is now treated as a doctrine of standing for private-plaintiff actions; in assessing the plaintiff’s standing under this doctrine, courts assume for the sake of argument that there is an antitrust violation.²⁸⁰

²⁷⁷ *Id.*

²⁷⁸ *Brunswick*, 429 U.S. at 487–89 (quoting *Zenith Radio Corp. v. Hazeltine Research Inc.*, 395 U.S. 100, 125 (1969)) (first and third emphases added).

²⁷⁹ See 2A AREEDA & HOVENKAMP, *supra* note 29, ¶ 337 (3d ed. 2007) (discussing scope of antitrust injury doctrine and cases).

²⁸⁰ See *id.*, ¶335f (“to test standing, assume a violation”).

Brunswick's most notable feature is its virtual disregard of the language of antitrust's statutory private action provision, which is in fact quite categorical. It speaks of "any person" who is injured in business or property "by reason of" an antitrust violation.²⁸¹ The statute itself says nothing about relating the nature of the injury to the competition-furthering goals of antitrust. Yet, today it is clear that the antitrust injury rule applies to equity actions just as much as to damages actions. Although antitrust's injunction provision requires only threatened harm of any type,²⁸² the Supreme Court has held that "[i]t would be anomalous . . . to read the Clayton Act to authorize a private plaintiff to secure an injunction against a threatened injury for which he would not be entitled to compensation if the injury actually occurred."²⁸³

In IP today, as in antitrust thirty years ago, the courts too often equate the issue of actionable injury with the issue of infringement. As a first step in their own errands into the wilderness, the intellectual property laws need to develop a conception of "IP injury," which would require not merely injury-in-fact or speculative injury, but rather demonstrable injury that is tied to the purpose for which the IP laws were passed in the first place. IP law should recognize harm only for uses that are likely to interfere with IP holders' decisions to create or distribute their works—that is, only for harms that are consistent with the constitutional mandate that the purpose of the patent and copyright systems is to further innovation.

Of course, in some sense, injury-in-fact always accompanies infringement: the IP holder can say that the defendant could have paid a license fee for the use, even if the defendant's use does not supplant sales of the copyrighted work or patented invention. After all, the defendant is free-riding on the IP holder's work, and the IP holder should at least be allowed to share in the benefits that the defendant gained from its use.

This argument reflects a view that free-riding is a wrong in and of itself, and that property rights in IP should internalize all of the benefits that the IP produces for others. The argument both overstates and mischaracterizes the nature of free-riding and the harm that it causes. Further, it fails to relate free-riding to the incentive to create. As Frischmann and Lemley have observed, "Spillovers—uncompensated benefits that one person's activity provides to another—are everywhere."²⁸⁴ It is both impossible and undesirable to compensate for all of them.²⁸⁵

²⁸¹ 15 U.S.C. §15.

²⁸² 15 U.S.C. §26 (giving a private injunction right to any person facing "threatened loss or damage by a violation of the antitrust laws").

²⁸³ *Cargill, Inc. v. Monfort of Colo., Inc.*, 479 U.S. 104, 112 (1986).

²⁸⁴ Brett M. Frischmann & Mark A. Lemley, *Spillovers*, 107 COLUM. L. REV. 257, 258 (2007) (providing examples).

²⁸⁵ *See id.* at 282–84 (concluding that "there is good economic evidence that greater innovation spillovers are

Spillovers are one of the true benefits of investment in all kinds of property rights, intellectual or otherwise.²⁸⁶ Sellers of butter benefit from the production of better bread, yet bread producers are not entitled to a percentage of butter sales. Nor is the author of a popular, copyrighted opera entitled to a cut of a firm's sales of opera glasses. Innovation has always benefited people other than the innovator, and the law has not attempted to capture all of those external benefits and return them.²⁸⁷ Quite the contrary, the IP Clause of the Constitution seeks to encourage innovation not for the benefit of the individual inventor, but rather for the benefit of society at large.²⁸⁸ Significant innovations tend to produce many valuable spillovers – that is why we encourage them.²⁸⁹

Indeed, innovators are themselves the beneficiaries of spillovers. For example, the development of patented or copyrighted software benefits greatly from a market that contains computers. The developer of a patented toaster benefits greatly from the fact that homes are equipped with electricity and that bakers make bread. But we do not require these inventors to compensate computer manufacturers, electric utilities or wholesale bakers. Or to view the issue from the other side, the developers of computers or builders of electric utilities will not be compensated for every benefit that their work confers on the makers of collateral and complementary products. Uncompensated spillovers go in both directions, and indeed, some IP rights become very valuable precisely because they sit atop the technology developed by others. For example, Microsoft Windows' value depends entirely on the existence of a computer hardware infrastructure that Microsoft itself did not develop.

associated with more, not less, innovation"); Julie E. Cohen, *Creativity and Culture in Copyright Theory*, 40 U.C. DAVIS L. REV. 1151 (2007); Brett M. Frischmann, *Spillovers Theory and Its Conceptual Boundaries*, WM. & MARY L. REV. (forthcoming 2009) (manuscript available at <http://ssrn.com/abstract=1357688>); see also Wendy J. Gordon, *Harmless Use*, 77 FORDHAM L. REV. 2411, 2419 (2009) (arguing that, taken to its logical conclusion, the prohibition against reaping where one has not sown "would be insane" because "in a civilized nation full of physical, technological, and cultural infrastructure, we reap from birth what others have sown").

²⁸⁶ See Harold A. Demsetz, *The Private Production of Public Goods*, 13 J.L. & ECON. 293 (1970); see also Mark A. Lemley, *Property, Intellectual Property, and Free Riding*, 83 TEX. L. REV. 1031 (2005).

²⁸⁷ See Frischmann & Lemley, *supra* note 284, at 259.

²⁸⁸ See, e.g., *Kendall v. Winsor*, 62 U.S. (21 How.) 322, 327–28 (1859) (“[T]he limited and temporary monopoly granted to inventors was never designed for their exclusive profit or advantage; the benefit to the public or community at large was another and doubtless the primary object in granting and securing that monopoly.”).

²⁸⁹ Frischmann & Lemley, *supra* note 284, at 284 (“[T]here is good economic evidence that greater innovation spillovers are associated with more, not less, innovation, which suggests that if anything we have too much IP protection and too few spillovers today.”); Richard A. Posner, *Do We Have Too Many Intellectual Property Rights?*, 9 MARQ. INTELL. PROP. L. REV. 173 (2005) (IP system recognizes too many rights, producing too little in uncompensated social benefit).

IP law in general has become much too obsessed with the notion that spilling over is itself a wrong for which compensation is due, or that the IP holder is entitled to compensation for every external benefit that its right produces, whether or not there is a reduction in the incentive to innovate. Spillovers are the “surplus” that innovation produces, just as the consumers’ surplus, or the difference between value and price, is the surplus that well-functioning markets produce. Requiring IP holders to be compensated for spillovers would be tantamount to requiring monopolists to be compensated up to the full value that customers place on the goods they sell. But antitrust is built on the bedrock principle that consumers are antitrust’s protected class; it is they, and not manufacturers, who are entitled to any surplus that increased competition products.²⁹⁰

IP law needs a similar approach that does not compensate for spillovers but requires compensation only where the defendant’s use harms the IP holder’s ex ante incentives to innovate. In many cases, the defendant’s use actually increases the IP holder’s incentives to innovate by increasing sales of the original work. Consider the British *Da Vinci Code* case. Dan Brown wrote the blockbuster novel *The Da Vinci Code*, a religious historical thriller involving a romantic relationship that allegedly existed between Jesus Christ and Mary Magdelene. Brown’s book borrowed heavily from a nonfiction book by Michael Baigent, Richard Leigh, and Henry Lincoln entitled *Holy Blood, Holy Grail*, which explored the Jesus/Mary Magdelene story. When *Da Vinci Code* was published, *Holy Blood*’s sales increased dramatically, restoring it briefly to the British best-seller list.²⁹¹ But that did not prevent its publisher from filing a copyright infringement suit in a British court.²⁹² The court eventually found no infringement because, while Brown apparently borrowed uncopyrightable facts and ideas, he borrowed only a trivial amount of copyrighted expression.²⁹³

Clearly, *The Da Vinci Code* is not a substitute for but a complement to *Holy Blood, Holy Grail*. When goods are complements, sales of one *increase* sales of the other—for example, cheaper butter increases bread sales and better software increases computer sales. Product complementarity increases rather than decreases the incentive to innovate. As a result, permitting

²⁹⁰ See Herbert Hovenkamp, *Antitrust’s Protected Classes*, 88 MICH. L. REV. 1 (1989).

²⁹¹ See the CBS News Story, which noted that sales of the plaintiff’s book had increased some 3500%. See James M. Klatell, Dramatic Day at ‘Da Vinci’ Trial, <http://www.cbsnews.com/stories/2006/03/07/entertainment/main1379394.shtml> (last visited July 30, 2009).

²⁹² See Richard Roth, Mining Da Vinci, http://www.cbsnews.com/stories/2006/03/10/listening_post/main1390534.shtml (last visited July 30, 2009). See also Bohannon, *supra* note 259, at 16; Bohannon, *supra* note 85, at 1029;

²⁹³ Baigent v. Random House Group Ltd., [2006] EWHC (Ch) 719 No. HC04C03092 (U.K.) (finding mainly copying of ideas rather than expression), *aff’d*, [2007] EWCA (Civ) 247 No. A3 2006/0971.

infringement actions against complements is inconsistent with the goals of copyright law unless the complementary work is one that the copyright owner was very likely to develop itself. While the authors of *Holy Blood, Holy Grail* might conceivably have contemplated a novel based on their nonfiction work, it certainly would not have been the novel that Dan Brown wrote. Ex ante, the prediction that someone will base a best-selling novel on the author's nonfiction historical book would almost certainly have increased rather than reduced that author's incentive to write it.

Spillovers become a problem only when they seriously threaten the incentive to invest in the first place. Spillovers that cause no foreseeable harm or that actually benefit the IP holder do not fall into this category. Under current IP infringement standards, about the only thing that we can say *ab initio* about an infringing work is that it is a spillover. Without analysis, we cannot conclude whether it harms, benefits, or has no impact on the value of an IP holder's rights. We certainly cannot say whether condemning it is worth the costs of doing so, including the loss of innovation on the part of others.

Conclusion: Ex Ante Harm and the Scope of IP Injury

IP owners have been enriched all these many years by per se damages provisions that simply presume harm and base damages on factors such as the infringer's profits or largely hypothetical lost licensing fees without any proof that the IP owner has been harmed in a way that is likely to diminish her incentive to innovate in the first place. As our experience in antitrust law has shown, one problem with per se rules of this sort is that they fail to make records about actual effects. For example, one of the reasons that we know so little about the economic consequences of resale price maintenance, or supplier fixing of the prices at which their goods can be resold, is that resale price maintenance was unlawful per se from 1911 until 2007.²⁹⁴ During that time, a plaintiff proved illegality by simply showing the agreement to maintain prices, and no one cared about anticompetitive effects.²⁹⁵ As a result, the litigation process rarely pursued them. The same is true in IP cases. For instance, in the *Wind Done Gone* case discussed earlier, the lower court's preliminary injunction, based in part on a presumption that the defendant's copying would cause irreparable harm, precluded the development of facts regarding the effects of the defendant's parody on the market for the plaintiff's work.

By contrast, requiring serious proof of harm in IP infringement cases would enable courts to develop a set of records that could provide a great deal of information about the kinds of infringements that actually diminish IP holders' investment expectations. In order to limit IP

²⁹⁴ See *supra* note 10.

²⁹⁵ At least, not until the Supreme Court imposed the "antitrust injury" doctrine on maximum resale price maintenance in *Atlantic Richfield v. USA Petroleum Co.*, 495 U.S. 328 (1990) (antitrust injury doctrine bars recovery from rivals claiming that prices imposed by competitor on its dealers were too low; it did not matter that at the time maximum resale price maintenance was unlawful per se).

protection to that which is necessary to encourage innovation, IP injury must measure incentives *ex ante* rather than *ex post*. Infringement harms innovation when it diminishes anticipated returns, and the only things that can be anticipated are those that are reasonably foreseeable.²⁹⁶ Indeed, if an act of IP infringement does no harm to the rights owner but benefits either the infringer or its customers and does not affect anyone else, then that act is a pure Pareto improvement—an economic ideal that is true of very few involuntary transactions. Condemning such an act would be inefficient legal policy even without regard to the lost incentives to innovate that it creates and the transaction costs of employing the legal system. The trick, of course, is defining the circumstances under which the IP owner is injured. An infringement in a market that the IP holder would never have entered anyway “injures” him only in the sense that he could have obtained a royalty but did not—that is, it fails to make him better off.

In both patent and copyright infringement, a serious harm requirement would entail remedies for “naked” infringement, or instances where acts of infringement clearly deprive the IP right holder of sales. It would also entail remedies for infringements in obvious markets where the infringer is not a competitor but it is clear that an innovator would rely on those royalties in deciding whether to create the work. This would include, for example, movie versions of copyrighted novels and short stories, translations, or lengthy published excerpts. On the other hand, it would suggest little or no protection for situations where any harm caused by the alleged infringement is merely speculative. If a particular use is unlikely to affect the IP holder’s decision to produce a work, then it should not be deemed infringing. Thus, there should ordinarily be no finding of infringement where the defendant produces a complementary work that increases sales of the protected work or in which the defendant uses the work for personal, non-commercial purposes for which ordinary people are likely unwilling to pay.

The injury requirement would also preclude finding infringement where the infringer’s use occupies a market that is remote from the IP holder’s original market. Here, a page of antitrust law is helpful, particularly the law of nascent or unestablished businesses as plaintiffs. Antitrust permits firms that are about to enter new markets to recover for antitrust violations, but only under strictly defined circumstances. The plaintiff who claims “I would have entered this market but for the defendant’s anticompetitive conduct” must typically show “intent and preparedness,” which generally requires an investment, a set of business plans, or other commitments to show that the defendant actually deprived the plaintiff of an opportunity that it had foreseen and to which it had made a substantial commitment.²⁹⁷

²⁹⁶ See Bohannon, *supra* note 85, at 988–89; Shyamkrishna Balganesh, *Foreseeability and Copyright Incentives*, 122 HARV.L.REV. 1569 (2009).

²⁹⁷ See 2A AREEDA & HOVENKAMP, *supra* note 29, ¶ 349 (3d ed. 2007); see, e.g., *Aviation Upgrade Techs., Inc. v. Boeing Co.*, 78 Fed. App’x 623 (9th Cir. 2003) (denying standing to nascent firm that had no experience in the market, no plant, no employees other than its principal, no financing, and did not have required FAA certification); *Ashley Creek Phosphate Co. v. Chevron USA*, 315 F.3d 1245 (10th Cir. 2003) (firm that owned some mineral leases but had not yet determined whether entry into phosphate production would be profitable, had not obtained financing, and had no

The unestablished business rules in antitrust are rules about standing, or entitlement to sue. Like the antitrust-injury doctrine, they were made in the face of a statute that proclaims that *anyone* who is injured in her business or property has an entitlement to damages.²⁹⁸ Further, since they are rules of standing, they assume that there was in fact an antitrust violation, just as the antitrust-injury doctrine does.²⁹⁹ These rules of standing are driven by the great ambiguity that attends any determination that exclusionary conduct actually violates the antitrust laws, about appraising the degree of harm in the absence of evidence of a clearly injured plaintiff, and our desire to create the right set of incentives for businesses to compete aggressively without worrying too much about harming rivals, particularly when harms to investment incentives are difficult to estimate.

Similar considerations apply in IP infringement cases. Although it is much easier to decide to license a particular use of IP than to build a factory, IP holders must prove more than speculative harm. We want innovators to have sufficient incentives to innovate, but not to ride roughshod over the public domain or the innovative ideas of others. Efficient incentives do not require capture of the value of every spillover.

It is no answer to say that IP owners “expect” that they will be compensated for every infringement, whether foreseeable or not. That rationale confuses *ex ante* and *ex post* expectations. *Ex post*, the farmer gets to keep the gold that was buried on his property. But *ex ante*, the possible presence of gold affects his willingness to purchase the land and the price he will pay only to the extent that he anticipates that the gold is there and he can place a probability and value on its presence. Innovators are presumably not willing to spend infinite dollars on a research project simply because the sum of all conceivable foreseeable and unforeseeable results might be infinitely large. Their willingness to invest, just as that of any other reasonable actor, is a function of what they can anticipate. To the extent the publisher of a novel can foresee a good market for a translation or movie rights, she will be willing to pay more. If the publication is a law book, she may foresee translation rights but probably not movie rights. Antitrust law has developed rules for determining the types of plaintiffs that have a sufficient, objectively-determined commitment to a certain market that they have not yet entered.³⁰⁰ On the accused infringer’s side, reliable predictions about what does and does not constitute an infringement will also greatly facilitate the incentive to make efficient investments. When a new venture was objectively foreseeable to the IP holder at the time the work was developed, it will also be to an

experience in production, lacked antitrust standing); *Bourns, Inc. v. Raychem Corp.*, 331 F.3d 704 (9th Cir. 2003) (similar).

²⁹⁸ See *supra* text accompanying note 275-277.

²⁹⁹ See 2A AREEDA & HOVENKAMP, *supra* note 29, ¶ 335f (3d ed. 2007).

³⁰⁰ *Id.* ¶349.

innovator who wishes to borrow from the IP holder's work. In that case, the innovator will know she must obtain a license.

An IP injury requirement would help to limit the scope of IP rights to the purpose for which they are granted. In addition, the IP injury requirement can help to provide adequate notice of IP claims. The best way to facilitate this set of values in an uncertain world is to give IP right holders the right to exclude in markets where they are actually operating, as well as markets in which they have made an actual initial investment or in which one can reasonably foresee investment as likely, looking from the time that a decision to commit resources to innovation is made.

Antitrust reform came about almost entirely through judicial rather than legislative initiatives. From the late 1970s and until today, the Supreme Court has acted to dismantle most of the expansionism that had occurred in the 1960s and earlier. The Supreme Court has been taking the lead in patent reform as well. What will happen in copyright is difficult to say. When the courts get the urge for reform, there is little to stop them—as in the case of antitrust, where they were inspired to produce rulings that were sensible but quite inconsistent with statutory language.³⁰¹ Whether reform impulses will drive the judiciary to such lengths in IP law is less certain. The statutes are more elaborate, more recent, and subject to more frequent Congressional intervention. On the other side, however, there is also considerable evidence that Congress is “stuck”—that it cannot muster the collective will to engage in serious reform, and then it resorts to interest groups to become its statutory drafters. For this reason, we think that the judicial system has a comparative advantage, and the imbalance in favor of the judiciary is even stronger in IP than it is in antitrust, because the extent of legislative capture is greater.

In that case, courts should keep in mind that the patent and copyright laws have explicit authorization in the Constitution—much more explicit than the general language of the Commerce Clause that enables the antitrust laws—and that this authorization expressly ties the IP rights created to the incentive to create. Thus, IP has a powerful guiding principle. It need only be used.

³⁰¹ In addition to *Brunswick*, discussed *supra* text accompanying notes 263–72, see also *Ill. Brick Co. v. Illinois*, 431 U.S. 720 (1977) (only direct purchasers may maintain action for overcharge damages), and *2A AREEDA & HOVENKAMP*, *supra* note 29, ¶ 341 (3d ed. 2007).