The Procompetitive Interest in Intellectual Property Law

Thomas F. Cotter*

Abstract: When government recognizes intellectual property (IP) rights, it is often viewed as sanctioning the existence of private "monopolies," in contrast to the general antimonopoly policy manifested in the antitrust laws. And yet, on occasion IP law itself condemns conduct on the part of IP owners--or excuses otherwise infringing activity on the part of IP defendants--expressly for the purpose of promoting competition, even though antitrust law (if one were to apply it at all under analogous circumstances) would not find anticompetitive harm absent a more thorough analysis of whether the antitrust defendant possesses power over a well-defined market. Salient examples include the misuse doctrines in patent and copyright law; some applications of merger and fair use in copyright; and trademark law's functionality doctrine. In this paper, I develop a theoretical explanation for this divergence between antitrust and IP. Specifically, I argue that in some limited contexts the expected social costs (including error costs) of ruling for IP defendants may be low in comparison with the expected anticompetitive harm from ruling for IP plaintiffs. As a result, it sometimes may be welfare-enhancing for IP courts to be less concerned than antitrust courts about the expected costs of "false positives," that is, cases wrongly decided against the party defending the allegedly anticompetitive conduct. I further contend, however, that such cases probably are more common in the copyright than in the patent law context, and that even in copyright contexts courts should be cautious about casually inferring anticompetitive harm; but that the analysis provides a rationale for a relatively expansive definition of trademark functionality.

* Professor of Law, Washington and Lee University School of Law. I thank Christina Bohannan, Jeff Harrison, Mark Lemley, Bill Page, Glen Robinson, Phil Weiser, David Zaring and participants at the Fifth Annual Intellectual Property Scholars Conference at Benjamin N. Cardozo School of Law, for their comments and criticism; Nicholas Frazier and Adina Pollan for research assistance; and Washington and Lee University School of Law, for a summer research grant. Special thanks to Professor Tom Vinson of Washington and Lee's Mathematics Department for assistance with the mathematical analysis. Any errors that remain are mine.

I. Introduction

Conventional wisdom holds that antitrust and intellectual property (IP) law seek to maximize social welfare in opposite ways, with antitrust law condemning monopoly to attain this goal and IP law granting temporary monopolies in order to achieve the same end.[1] This "wisdom" is an oversimplification insofar as (1) contemporary antitrust enforcers, in recognition of the fact that IP right typically do not give rise to economically meaningfully monopolies,[2] approach IP-related conduct and transactions with much less inherent suspicion than often was the case in years past;[3] and (2) antitrust law by itself does not condemn the mere monopoly possession of monopoly power, but rather certain exercises of or efforts to obtain it.[4] Nevertheless, there is a kernel of truth to the conventional characterization, to the extent that IP rights *sometimes* enable the exercise of a degree of market power,[5] and *sometimes* even permit IP owners to engage in conduct

---

[1] *See, e.g.*, Michael Carrier, *Unraveling the Patent-Antitrust Paradox*, 150 U. PENN. L. REV. 761, 762-63 (2002); Louis Kaplow, *The Patent-Antitrust Intersection: A Reappraisal*, 97 HARV. L. REV. 1815, 1817, 1818 n.10 (1984).

[2] *See See* U.S. DEP'T OF JUSTICE & FEDERAL TRADE COMM'N, ANTITRUST GUIDELINES FOR THE LICENSING OF INTELLECTUAL PROPERTY § 2.1 (Apr. 6, 1995) [hereinafter, "IP GUIDELINES"]; *see also* 1 HERBERT HOVENKAMP ET AL., IP AND ANTITRUST § 4.2, at 4-9 to -10 (2002); WILLIAM M. LANDES & RICHARD A. POSNER, THE ECONOMIC STRUCTURE OF INTELLECTUAL PROPERTY LAW 374 (2003). *Cf.* Ariel Katz, *Intellectual Property, Antitrust, and the Presumption of Market Power: Making Sense of an Alleged Nonsense*, Univ. of Toronto Res. Paper No. 02-05, *available at* http://ssrn.com/abstract=702462 (arguing that commercially successful IP rights often do give rise to market power).

[3] *See* IP GUIDELINES, *supra* note 2, § 2.2 (stating that the "Agencies will not presume that a patent, copyright, or trade secret necessarily confers market power upon its owner," and that "there will often be sufficient actual or potential close substitutes" for an IP-protected "product, process, or work to prevent the exercise of market power"). In the view of the Department of Justice (DOJ) and the Federal Trade Commission (FTC), "[i]ntellectual property is . . . neither particularly free from scrutiny under the antitrust laws, nor particularly suspect under them." *Id.* § 2.1. This approach contrasts with the widespread suspicion concerning the anticompetitive nature of IP rights throughout much of the twentieth century. *See, e.g.*, Willard K. Tom & Joshua A. Newberg, *Antitrust and Intellectual Property: From Separate Spheres to Unified Field*, 66 ANTITRUST L.J. 167, 178-84 (1997).

[4] *See infra* notes __ and accompanying text.

[5] Market power is "the ability profitably to maintain prices above, or output below, competitive levels for a significant period of time." IP GUIDELINES, *supra* note __, § 2.2; *see also* U.S. DEP'T OF JUSTICE &

that would be unlawful absent the IP right.[6] Though subject to exaggeration, the recognition and enforcement of IP rights might be seen as an exception to the anti-monopoly policy embodied in the antitrust laws.

There is another side of the coin, however, that is less frequently commented upon.[7] On occasion, IP law condemns conduct on the part of IP owners--or excuses otherwise infringing activity on the part of IP defendants--for the express purpose of promoting competition, even though antitrust law (if it were to apply at all) typically would not condemn similar conduct on the part of the IP owner (or require the IP defendant be given free access), absent thorough analysis of (1) the markets within which the parties compete, and (2) whether the IP owner possesses market power.[8] To illustrate, consider these examples of what I will refer to as IP law's "procompetitive doctrines":[9]

FEDERAL TRADE COMM'N, ANTITRUST HORIZONTAL MERGER GUIDELINES § 0.1 (rev. ed. Apr. 8, 1997) [hereinafter, the "HORIZONTAL MERGER GUIDELINES"] (same); U.S. DEP'T OF JUSTICE & FEDERAL TRADE COMM'N, ANTITRUST GUIDELINES FOR COLLABORATIONS AMONG COMPETITORS § 3.3 n.30 (Apr. 2000) [hereinafter, the "COLLABORATION GUIDELINES"]. The DOJ and the FTC define a market as "a product or group of products and a geographic area in which it is produced or sold such that a hypothetical profit-maximizing firm . . . that was the only present and future producer or seller of those products in that area likely would impose at least a 'small but significant and nontransitory' increase in price . . . ." HORIZONTAL MERGER GUIDELINES, *supra*, § 1.0.

[6] For example, the Patent Act permits patent owners to engage in horizontal territorial divisions with their licensees, *see* 35 U.S.C. § 261 (2000), a practice that otherwise would constitute a per se violation of the Sherman Act, *see* Palmer v. BRG, Inc., 498 U.S. 46, 48-50 (1990) (per curiam).

[7] There has been *some* previous commentary on these matters. *See, e.g.*, Robert G. Bone, *Enforcement Costs and Trademark Puzzles*, 90 VA. L. REV. 2099, 2174-81 (2004); Anna F. Kingsbury, *Market Definition in Intellectual Property Law: Should Intellectual Property Courts Use an Antitrust Approach to Market Definition?*, 8 MARQUETTE INTELL. PROP. L. REV. 63 (2004); Bone's approach is closest to mine, but it is limited to the trademark context.

[8] Antitrust law itself does not always require proof of harm to a defined market, *see infra* text accompanying note __, but it would be unlikely to condemn the type of conduct I discuss in this paper in the absence of such proof.

[9] Some other IP doctrines—for example, copyright's *scenes à faire* and functionality doctrines, the latter of which differs from trademark's functionality doctrine in some important ways--also can be thought of as embodying a balance between incentives, on the one hand, and competitive need, on the other, in some

- In patent and copyright cases, courts sometimes apply the *misuse* doctrine to render the IP plaintiff's patent or copyright unenforceable until the plaintiff "purges" that misuse.[10] Courts continue to state the principle that certain practices on the part of patent and copyright owners are sufficiently anticompetitive to constitute misuse, even though the same conduct would result in antitrust liability, if at all, only upon proof of additional elements including market power.[11]

- In copyright law, courts sometimes excuse defendants from liability by applying the *merger* or *fair use* doctrines in contexts in which the copyright owner's exercise of exclusive rights, often in some aspect of its computer software, might impede competition in a related market.[12] Courts do not require the defendant to define what that market is, however, or to quantify the harm, in any way comparable to the plaintiff's burden in analogous antitrust litigation.

- In trademark law, the *functionality* doctrine permits defendants to copy a trademark owner's distinctive product design, notwithstanding the potential for consumer confusion, if *inter alia* the copying is justified by competitive need.[13] As in the above two examples, however, courts rarely articulate what they mean by "competitive need," and do not apply antitrust standards to determine whether according exclusive rights would impede competition in any well-defined market.

Two possible justifications for IP law's comparatively loose conceptualization of "markets" and "competition" are what might be referred to as the "beyond the scope" and "adjudication cost" rationales. The "beyond the scope" rationale posits that courts correctly excuse IP defendants from liability when not doing so would enable the IP

---

instances. *See* Michael J. Meurer, *Vertical Restraints and Intellectual Property Law: Beyond Antitrust*, 87 MINN. L. REV. 1871, 1876, 1910 n.232 (2003). For present purposes, however, I confine my discussion to what appear to me to be the principal doctrines with respect to which courts explicitly consider competitive need.

[10] *See, e.g.*, Va. Panel Corp. v. MAC Panel Co., 133 F.3d 860, 869 (Fed. Cir. 1997); Lasercomb Am., Inc. v. Reynolds, 911 F.2d 970, 973-77 (4th Cir. 1990).

[11] *See, e.g.*, Alcatel U.S.A., Inc. v. DGI Techs., Inc., 166 F.3d 772, 793 (5th Cir. 1999); Senza-Gel Corp. v. Seiffhart, 803 F.2d 661, 668 (Fed. Cir. 1986).

[12] *See, e.g.*, Sega Enters. Ltd. v. Accolade, Inc., 977 F.2d 1510, 1523-24 (9th Cir. 1993); (fair use); Herbert Rosenthal Jewelry Corp. v. Kalpakian, 446 F.2d 738, 742 (9th Cir. 1971) (merger). Fair use also applies in several other settings for reasons not directly related to competitive need. *See infra* notes __ and accompanying text. My focus here is on the competitive need line of fair use cases only.

[13] *See, e.g.*, TrafFix Devices, Inc. v. Marketing Displays, Inc., 532 U.S. 23, 32-33 (2001) (product feature trade dress is functional, and hence not subject to trademark protection, if, *inter alia*, the exclusive use of the feature "would put competitors at a significant non-reputation-related disadvantage").

owner to control subject matter that the IP laws themselves, presumably for reasons bearing some relation to competitive need among other considerations, relegate to the public domain ("beyond the scope" of the IP owner's rights). I will argue that this justification is useful up to a point, but that ultimately it depends upon a preexisting baseline—a principle that IP rights extend only up to Point X and no farther—which is only sometimes readily discernible.[14] Everyone may agree, for example, that copyright does not subsist in ideas; but exactly what distinguishes an "idea" from "expression" is often unclear, and in many instances the distinction must be made in light of its predictable consequences. Put another way, formalistic application of IP's procompetitive doctrines to prevent the expansion of IP rights beyond their scope might undermine consideration of how best to interpret that scope so as to attain an optimal accommodation of monopoly incentives with competition. Alternatively, under an "adjudication cost" rationale, one might justify IP law's less rigorous approach to competitive need on the ground that it would be unduly expensive in some contexts to prove competitive harm in the same way that antitrust often requires. Potentially turning every merger doctrine or trade dress case into a miniature antitrust dispute might seem extravagant.[15] Like the beyond the scope argument, however, this argument also rests upon a baseline assumption, in this instance that the social benefits of protecting the IP rights at issue are not sufficiently large as to warrant the investment of substantial resources into more accurately evaluating competitive need. The assumption may be

[14] For similar observations, see Kaplow, *supra* note __, at 1848-49; Glen O. Robinson, *Personal Property Servitudes*, 71 U. CHI. L. REV. 1449, 1469 (2004); Note, *Clarifying the Copyright Misuse Defense: The Role of Antitrust Standards and First Amendment Values*, 104 HARV. L. REV. 1289, 1295 (1991).

[15] *See, e.g.*, Maureen O'Rourke, *Drawing the Boundary Between Copyright and Contract: Copyright Preemption of Software License Terms*, 45 DUKE L.J. 479, 550-51 (1995). Presumably many litigants would abandon these issues if courts did require such extensive proof.

correct, but it would be useful to articulate why, when the costs of greater certainty are prohibitive, some cases involving speculative competitive harms should be resolved in favor of IP defendants rather than IP plaintiffs.

Notwithstanding the preceding reservations, my thesis is that IP law on occasion may increase social welfare by promoting competition in ways that antitrust law does not address, and based on evidence that would be insufficient in an antitrust context. Context, in other words, is crucial; although forced sharing to attain optimal competition might seem unwarranted in most antitrust contexts, absent clear proof of market harm, it might constitute good IP policy even in the presence of ambiguous evidence. A major premise of the argument is that both antitrust and IP law can be viewed as seeking to minimize the expected cost of (1) "false positives," that is, cases incorrectly[16] decided in favor of the IP defendant or the antitrust plaintiff; (2) "false negatives," cases incorrectly decided in favor of the IP plaintiff or antitrust defendant; and (3) adjudication. In seeking to attain this goal, however, antitrust and IP may diverge insofar as (1) certain types of "false positive" harms that are likely to be present in the antitrust context may, on occasion, be nonexistent in the IP context; and (2) certain types of "false negative" harms that are not cognizable in the antitrust context *are* worth taking into account in the IP context. Apropos of the latter, I will argue that IP law sometimes should be less tolerant

[16] "Incorrectly" means that a decision to the contrary would have produced more social benefits than social costs. A case decided incorrectly in favor of an antitrust plaintiff or IP defendant (a "false positive") therefore is one in which social welfare would have been better served by ruling in favor of the antitrust defendant or IP plaintiff. A case decided incorrectly in favor of an antitrust defendant or IP plaintiff (a "false negative") is one in which social welfare would have been better served by ruling in favor of the antitrust plaintiff or IP defendant. For discussion of how antitrust rules can be viewed as efforts to minimize the expected cost of false positives, false negatives, and adjudication costs, see, e.g., C. Frederick Beckner III & Steven C. Salop, *Decision Theory and Antitrust* Rules, 67 ANTITRUST L.J. 41 (1999); John E. Lopatka & William H. Page, *Monopolization, Innovation, and Consumer Welfare*, 69 GEO. WASH. L. REV. 367, 387 (2001). For application of this type of analysis in the IP literature, see, e.g., Bone, *supra* note __, at 2125-81; Katz, *supra* note __, at 48-54.

than antitrust of (1) harms stemming from the exercise of small-scale, transitory, or localized market power; (2) harms that are individually of little competitive significance, but which in the aggregate threaten to reduce social welfare; and (3) threats to dynamic efficiency stemming from the reduction of competition in so-called "innovation markets." In addition, although it often may be good antitrust policy to incur substantial adjudication costs to define the market and ascertain anticompetitive consequences, in the IP context it may be welfare-enhancing to avoid these costs if the stakes are sufficiently low; or if one can be reasonably confident absent such analysis that the expected costs (including error costs) of ruling for the plaintiff outweigh the expected costs of ruling for the defense.

Part II begins with a brief overview of the role of market definition in antitrust law, and of the restrictions that antitrust places upon the exercise of IP rights. Part III then shows how courts sometimes apply IP doctrine to promote competition in ways that diverge from antitrust standards, focusing on the examples referred to above from each of the three main branches of IP law: misuse, fair use and merger, and functionality. Part IV fleshes out the analysis of false positives and false negatives, and argues that certain harms should be of more interest in the IP than the antitrust context. Part V applies this analysis to the misuse, fair use, merger, and functionality doctrines. I conclude that, notwithstanding the theoretical case for a more aggressive competition policy in some IP contexts, courts should be wary of applying the misuse doctrine in all but the most exceptional circumstances; but that a less rigorous analysis of competitive harm may make sense in some contexts involving fair use, merger, and (especially) functionality.

II. The Antitrust Context

The dominant view today among antitrust courts and scholars is that antitrust law is a tool for promoting social welfare by deterring practices and transactions that tend to increase market power.[17] Economic theory predicts that the monopolist maximizes profits by charging a higher price, and producing less output, than would a competitive firm, and that the monopolist's gain from so doing is less than consumers' loss (the "deadweight loss").[18] As a first approximation, then, measures that deter monopoly increase social wealth; but this principle is subject to a number of qualifications. The first is that the mere *possession* of monopoly power, as opposed to its willful acquisition or maintenance, is not an antitrust offense.[19] Second, antitrust policymakers have come to recognize that the prospect of attaining market power may encourage innovation; indeed, this is the principal rationale for granting IP rights to innovators.[20] Partially in response to this insight, antitrust law today concerns itself not only with the static inefficiency defined by deadweight loss, but also with improving consumer welfare over

[17] *See* John E. Lopatka & William H. Page, *Economic Authority and the Limits of Expertise in Antitrust Cases*, 90 CORNELL L. REV. 633-34, 637-38 (2005); Steven L. Schwarcz, *Private Ordering*, 97 NW. U. L. REV. 319, 332 (2002).

[18] *See* JEAN TIROLE, THE THEORY OF INDUSTRIAL ORGANIZATION § 1.1.1.2, AT 67 (1988).

[19] *See* Verizon Comm. Inc. v. Law Office of Curtis V. Trinko, LLP, 540 U.S. 398, 407 (2004).

[20] *See, e.g.*, F.M. Scherer, *The Innovation Lottery*, *in* EXPANDING THE BOUNDARIES OF INTELLECTUAL PROPERTY: INNOVATION POLICY FOR THE KNOWLEDGE SOCIETY 3, 19-21 (Rochelle Cooper Dreyfuss et al. eds., 2001). One influential view of innovation, first advanced by the economist Joseph Schumpeter in the 1940s, and still debated among innovation theorists, is that the gains from innovation outweigh the relatively transitory losses from monopoly. *See* JOSEPH A. SCHUMPETER, CAPITALISM, SOCIALISM AND DEMOCRACY 84-85, 100-03 (1976) [1942]. On this logic, measures to encourage innovation should dominate over the comparatively trivial concerns of antitrust; the gains from innovation are permanent, whereas no monopolist can maintain its position forever. Note also that trademark law, unlike patent and copyright, is designed principally to lower consumer search costs, by according exclusive rights to source identifiers; but in so doing, it provides an incentive for producers to invest in quality control. *See infra* text accompanying note __.

the long run (so-called "dynamic" efficiency).[21] Most importantly, perhaps, antitrust standards must accommodate the reality that decisionmaking is susceptible to error, and that efforts to reduce the incidence of error can be costly. Antitrust policy therefore should take into account both the risk of error (including false positives, i.e., that courts will wrongfully condemn as inefficient conduct that increases social welfare, and false negatives, i.e., that they will wrongfully permit conduct that reduces welfare), as well as judicial and other administrative costs.[22] Applying different standards of antitrust scrutiny to different conduct can be seen as an effort to minimize the sum of these costs.

To illustrate, courts have interpreted section 1 of the Sherman Act, which condemns contracts, combinations, and conspiracies in restraint of trade,[23] to render per se illegal practices including horizontal price-fixing, output restrictions, and geographic divisions, as well as vertical minimum retail price maintenance.[24] The stated rationale behind per se condemnation is that practices like these are so likely to impede competition that a detailed inquiry into their effects in particular cases would not justify the additional costs (including both the additional administrative and judicial costs such an inquiry would entail, and the greater uncertainty borne by potential litigants).[25] Other

---

[21] *See, e.g.*, Ronald W. Davis, *Innovation Markets and Merger Enforcement: Current Enforcement in Perspective*, 71 ANTITRUST L.J. 677, 682-83 (2003).

[22] *See* Beckner & Salop, *supra* note __, at 44-46; Lopatka & Page, *supra* note __, at 387. The expected cost of an erroneous decision is a function of the frequency of error, which can sometimes be reduced by investing more in adjudication costs, and the magnitude of the harm flowing from the error. *See* Bone, *supra* note __, at 2124.

[23] 15 U.S.C. § 1 (2000).

[24] *See, e.g.*, Palmer v. BRG, Inc., 498 U.S. 46, 48-50 (1990) (per curiam) (geographic divisions); NCAA v. Bd. of Regents, 468 U.S. 85, 99 (1984) (output restrictions); Arizona v. Maricopa Cty. Med. Soc'y, 457 U.S. 332, 344-55 (1982) (horizontal price fixing); Euromodas, Inc. v. Zanella, Ltd., 368 F.3d 11, 16 (1st Cir. 2004) (vertical minimum resale price maintenance).

[25] *See* Ragsdale v. Wolverine World Wide, Inc., 535 U.S. 81, 93 (2002).

types of arrangements, however, which courts view as offering plausible procompetitive benefits, are analyzed under the rule of reason, under which "the factfinder weighs all of the circumstances of a case in deciding whether a restrictive practice should be prohibited as imposing an unreasonable restraint on competition."[26] In practice, courts apply something of a sliding scale to determine how detailed an inquiry is necessary, given the likely effects of the arrangement in question upon competition.[27] As a result, often the most important issue is how to characterize a practice or transaction. Characterizing a transaction as horizontal price-fixing, for example, invokes per se illegality, whereas a different characterization, if plausible, may result in no assessment of liability.[28]

In cases involving per se illegality, defining the market is often unnecessary, because the practice at issue would make no sense unless it had the effect of monopolizing *some* product or geographic market.[29] For example, naked price fixing among competitors typically makes no economic sense unless the competitors comprise a

---

[26] Continental T.V., Inc. v. GTE Sylvania Inc., 433 U.S. 36, 49 (1977). *See, e.g.*, State Oil Co. v. Khan, 522 U.S. 3, 8-22 (1997) (holding that vertical maximum resale price maintenance is subject to the rule of reason); Business Elecs. Corp. v. Sharp, 485 U.S. 717, 726-36 (1988) (holding that vertical nonprice restraints are subject to the rule of reason); *Northwest Wholesale Stationers*, 472 U.S. at 295 (holding that horizontal refusals to deal are subject to the rule of reason, absent proof that the defendants possess market power); Jefferson Parish Hosp. Dist. No. 2 v. Hyde, 466 U.S. 2, 29-30 (1984) (holding that tying is subject to the rule of reason, absent proof that, *inter alia*, the seller possesses market power).

[27] *See* California Dental Ass'n v. FTC, 526 U.S. 756, 780-81 (1999); PolyGram Holding, Inc., FTC Docket No. 9298, at 13-35 (July 24, 2003), *available at* http://www.ftc.gov/os/2003/07/polygramopinion.pdf, *pet'n for review denied*, 416 F.3d 29 (D.C. Cir. 2005).

[28] An example is *Broadcast Music, Inc. v. Columbia Broadcasting System, Inc.*, 441 U.S. 1 (1979), in which the Court declined to characterize BMI's royalty scheme as price fixing, noting its likely efficiency in minimizing transaction costs. *See id.* at 8-9, 22.

[29] *See* HERBERT HOVENKAMP, FEDERAL ANTITRUST POLICY: THE LAW OF COMPETITION AND ITS PRACTICE § 5.6b, at 254 (2d ed. 1999).

dominant share of some discrete market.[30] In most other instances, however, courts will not presume anticompetitive effects but rather will require the antitrust plaintiff to prove them. Even then, formal market definition may not always be necessary. In *FTC v. Indiana Federation of Dentists*,[31] for example, the Supreme Court held that an agreement on the part of the defendants to withhold dental x-rays from insurance companies, while not per se illegal, could be condemned under a truncated version of the rule of reason after finding the proffered procompetitive justifications for the rule wanting.[32] Absent proof of "actual detrimental effects," however, such as higher prices or lower output or quality, once an agreement evades per se condemnation, the antitrust plaintiff must support its claim with evidence of the effect of the targeted transaction or practice upon competition in some well-defined market.[33]

The process of defining markets can be laborious. As a general matter, the courts and agencies define a market as "a product or group of products and a geographic area in which it is produced or sold such that a hypothetical profit-maximizing firm . . . that was the only present and future producer or seller of those products in that area likely would

---

[30] In a competitive market, price fixing *below* cost reduces the price fixers' profits below the break-even point, while price fixing *above* cost will induce consumers to buy from the price-fixers' non-colluding rivals. Logic therefore suggests that firms engaging in price fixing must control a substantial portion of a market, and that forgoing proof of market power therefore conserves judicial and party resources. *See* FTC v. Superior Ct. Trial Lawyers Ass'n, 493 U.S. 411, 430-31 (1990).

[31] 476 U.S. 447 (1986).

[32] *See id.* at 459-61. The Court did not require formal proof of the relevant market, reasoning that "since the purpose of the inquiries into market definition and market power is to determine whether an arrangement has the potential for genuine adverse effects on competition, 'proof of actual detrimental effects, such as a reduction of output,' can obviate the need for an inquiry into market power, which is but a 'surrogate for detrimental effects.'" *See id.* at 460-61 (quoting 7 PHILLIP E. AREEDA, ANTITRUST LAW ¶ 1511, at 424 (1986)).

[33] *See, e.g.*, Copperweld Corp. v. Independence Tube Corp., 467 U.S. 752, 768 (1984); COLLABORATION GUIDELINES, *supra* note __, § 3.3, at 11. Plaintiffs asserting other types of antitrust claims also must properly define the market. *See, e.g.*, Eastman Kodak Co. v. Image Tech. Servs., Inc., 504 U.S. 451, 481-82 (1992) (Sherman Act § 2); HORIZONTAL MERGER GUIDELINES, *supra* note __, § 1.0 (Clayton Act § 7).

impose at least a 'small but significant and nontransitory' increase in price . . . ."[34] In a competitive market a firm that imposes a "'small but significant and nontransitory' increase in price" would lose money, as its customers would switch to other suppliers of that product, or to other substitute products. Thus, in analyzing whether Product A occupies a discrete market, one must consider not only whether there are other sellers of Product A, but also whether Product A is interchangeable with other existing or potential products.[35] Ideally, one should consider the cross-elasticity of demand between Product A and other relevant products[36]--i.e., whether consumers would view Products B, C, and so on as adequate substitutes for A,[37] if the firm were increase the price of A above the competitive level[38]—and also the cross-elasticity of supply[39] (whether firms that manufacture products M, N, and so on would respond to an increase in the price of A by switching some of their production to A, or to other products that consumers view as adequate substitutes for A).[40] Due to the difficulty of estimating elasticities, however, enforcers often consider proxies for demand and supply-side substitutability.[41]

---

[34] *Id.* § 1.0, at 4.

[35] *See, e.g.*, Eastman Kodak Co. v. Image Tech. Servs., Inc., 504 U.S. 451, 482 (1992).

[36] *See* United States v. Grinnell Corp., 384 U.S. 563, 592-93 (1966); Worldwide Basketball & Sport Tours, Inc. v. Nat'l Collegiate Athletic Ass'n, 388 F.3d 955, 962 (6th Cir. 2005).

[37] The extent to which two products are substitutes can be measured by their cross-elasticity of demand. *See generally* ROGER D. BLAIR & DAVID L. KASERMAN, ANTITRUST ECONOMICS 108-10 (1983).

[38] Considering whether a firm could raise price above *the competitive level* (i.e., marginal cost) is an important qualification. Because a small upward increase from the monopoly price therefore probably *would* induce consumers to buy substitute products, using the monopoly price as a baseline defines the market too broadly. *See* RICHARD A. POSNER, ANTITRUST LAW 150 (2d ed. 2001).

[39] *See, e.g.*, AD/SAT v. Associated Press, 181 F.3d 216, 227 (2d Cir. 1999).

[40] *See* HORIZONTAL MERGER GUIDELINES, *supra* note __, § 1.320; POSNER, *supra* note __, at 150.

[41] *See* HORIZONTAL MERGER GUIDELINES, *supra* note __, § 1.11, at 6-7; *see also Rothery*, 792 F.2d 210, 218 & n.4.

The Guidelines state the agency should continue the process of defining the market in the following manner:

> If, in response to the [hypothetical] price increase, the reduction in sales of the product would be large enough that a hypothetical monopolist would not find it profitable to impose such an increase in price, then the Agency will add to the product group the product that is the next-best substitute for the merging firm's product. . . .
>
> The price increase question is then asked for a hypothetical monopolist controlling the expanded product group. . . . The process will continue until a group of products is identified such that a hypothetical monopolist over that group of products would profitably impose at least a 'small but significant and nontransitory' increase . . . . The Agency will consider the relevant product market to be the smallest group of products that satisfies this test.[42]

As a general matter, a "price increase of five percent lasting for the foreseeable future" constitutes a "small but significant and nontransitory" increase, although "the Agency at times may use a price increase that is larger or smaller than five percent."[43] Although the preceding analysis was developed specifically for use in evaluating mergers, courts apply similar criteria for defining the market in § 1 and § 2 cases.

Although this is not the place for an exhaustive treatment of the subject, the exercise of IP rights can add further complexities to antitrust analysis. If the IP laws are to fulfill their purpose of stimulating creativity and other social benefits by offering the lure of potential monopoly profits, antitrust policy must accommodate the IP owner's exercise of market power up to some point. Even so, it is clear that some transactions involving IP can be subject to per se condemnation. Tying arrangements, for example, remain per se unlawful (though subject to many qualifications), and they are not exempt

---

[42] *See* HORIZONTAL MERGER GUIDELINES, *supra* note __, § 1.11, at 6.

[43] *Id.*

from antitrust scrutiny merely because the antitrust defendant conditions the sale or license of its IP, as opposed to some other good, on the purchase or license of another product.[44] Most other IP-related transactions, however, are evaluated under the rule of reason. IP licenses, for example, can embody a wide variety of restraints, including a commitment on the part of the licensee to confine its use to a specific territory, or to limit its use to certain purposes (a "field of use" restriction); or to grant back to the IP owner a license to use any IP that the licensee itself derives from the use of the owner's IP.[45] Alternatively, IP owners may agree to license one another's patents (a "patent pool"),[46] ostensibly to facilitate the parties' use of the patents and to reduce transaction costs. Agreements such as these, particularly if they are nonexclusive, are also likely to pass muster absent proof that the restraints (1) have specific anticompetitive effects and (2) are not reasonably necessary to achieve procompetitive efficiencies, such as the reduction of transaction costs and the efficient development and use of new technologies.[47]

As for monopolization claims, the law is clear that enforcing IP rights that have been acquired by fraud, or asserting "sham" infringement claims, can violate Sherman

[44] *See* Independent Ink, Inc. v. Illinois Tool Works, Inc., 396 F.3d 1342, 1346-52 (Fed. Cir. 2005), *cert. granted*, 125 S. Ct. 2937 (2005). Similarly, naked price-fixing is per se illegal, even if the "product" the price of which is being fixed is a patent or copyright. *See* IP GUIDELINES, *supra* note __, § 5.1, ex. 9; *id.* § 5.2. On the other hand, although a horizontal territorial division is usually a per se offense, an agreement conferring upon a patent licensee the exclusive right to market a patented product within a given territory is in general lawful. *See supra* note __.

[45] *See* IP GUIDELINES, *supra* note __, § 2.3, at 5; *id.* §§ 5.5, 5.6.

[46] *See* Robert P. Merges, *Of Property Rules, Coase, and Intellectual Property*, 94 COLUM. L. REV. 2655, 2662 n.27 (1994).

[47] *See* IP GUIDELINES, *supra* note __, §§ 4.2, 5.5, 5.6.

Act § 2.[48] Beyond these examples, however, there is some uncertainty concerning just how far antitrust should intrude into the domain of IP. Section 2 claims directed against a unilateral refusal to license one's IP, for example, present great difficulties, because the ability to decide if and when to license is often viewed as an essential aspect of IP rights. A few courts therefore have held or implied that a refusal to license IP can never constitute a § 2 violation, other than in cases involving fraud in the procurement or sham petitioning.[49] Others have held that a refusal to license IP is presumptively valid, but that the antitrust plaintiff can rebut the presumption by showing that the refusal was a pretext.[50] A considerable scholarly literature debates these competing views.[51]

The resolution of these issues is beyond the scope of this article; suffice to say that, under even the approaches that are more friendly to antitrust plaintiffs, antitrust law normally defers to the IP's owner refusal to deal with prospective licensees. Therein lies an apparent paradox. Antitrust law rarely condemns unilateral refusals to share property,

---

[48] *See* Professional Real Estate Investors, Inc. v. Columbia Pictures Indus., 508 U.S. 49, 60-61 (1993); Walker Process Equip., Inc. v. Food Mach. & Chem. Corp., 382 U.S. 172, 177-78 (1965); Nobelpharma AB v. Implant Innovations, Inc., 141 F.3d 1059, 1068-71 (Fed. Cir. 1998).

[49] *See* In re Independent Serv. Orgs. Antitrust Litig., 203 F.3d 1322, 1325-28 (Fed. Cir. 2000); Intergraph Corp. v. Intel Corp., 195 F.3d 1346, 1356-62 (Fed. Cir. 1999).

[50] *See* United States v. Microsoft Corp., 253 F.3d 34, 63-65 (D.C. Cir. 2001); Data Gen. Corp. v. Grumman Sys. Support Corp., 36 F.3d 1147, 1187 & n.64 (1st Cir. 1994); Image Tech. Servs. v. Eastman Kodak Co., 125 F.3d 1195, 1218-20 (9th Cir. 1997).

[51] *See, e.g.*, 3 PHILLIP E. AREEDA & HERBERT HOVENKAMP, ANTITRUST LAW ¶ 709b (2d ed. 2002); Joseph P. Bauer, *Refusals to Deal with Competitors by Owners of Patents and Copyrights: Reflections on the* Image Technical *and* Xerox *Decisions*, Notre Dame Law School Legal Studies Research Paper No. 05-12, *available at* http://ssrn.com/abstract=743365; Michelle M. Burtis & Bruce H. Kobayashi, *Why an Original Can Be Better Than a Copy: Intellectual Property, the Antitrust Refusal to Deal, and ISO Antitrust Litigation*, 9 S. CT. ECON. REV. 143 (2001); *see also* Glen O. Robinson, *On Refusing to Deal with Rivals*, 87 CORNELL L. REV. 1177, 1210 (2002) (arguing that antitrust's treatment of the refusal to license IP rights should not differ from its treatment of property rights generally). Robinson's view is consistent with Patent Act § 211, which states that "[n]othing in this chaper shall be deemed to convey to any person immunity from civil or criminal liability, or to create any defenses to actions, under any antitrust law." 35 U.S.C. § 211. *But see id.* § 271(d)(4) (stating that "[n]o patent owner otherwise entitled to relief . . . shall be . . . deemed guilty of misuse or illegal extension of the patent right by reason of his having . . . refused to license . . . any rights to the patent").

even more rarely condemns unilateral refusals to share *intellectual* property, and in any event requires proof of market definition and market power before undertaking to impose forced sharing. And yet IP law itself, *for the avowed purpose of promoting competition*, is sometimes much less generous to IP owners, granting IP users access without requiring comparable proof of market definition and market power. This paradox is not insoluble, but it does require fresh analysis of, among other things, the costs of error in IP litigation.

III. IP's Procompetitive Doctrines

Despite the fact that all of the IP doctrines referred to above are said to promote competition, courts almost never define the relevant markets within which the parties compete, or articulate why a particular result is necessary for effective competition. This omission does not necessarily mean that the courts always reach the wrong results, or that good reasons do not exist for IP's comparatively relaxed approach to these issues; but it does suggest a disconnect between the prevailing modes of analysis within these two bodies of law. This Part discusses IP law's procompetitive defenses—specifically, the misuse defense in patent and copyright law; the merger and fair use doctrines in copyright; and the functionality doctrine in trademark law--and shows how the courts' application of these doctrines contrasts with the approach that is often taken in antitrust.

A. Misuse

The misuse doctrine in patent and copyright law evolved from a series of early twentieth century Supreme Court decisions, all of which involved questions of whether patent owners could enforce their patents against persons who disregarded notices, affixed to patented products, stating that use of the product was subject to restrictions.

For example, in *Motion Picture Patents Co. v. Universal Film Manufacturing Co.*,[52] the Court held that a defendant who had lawfully acquired patented equipment from a middleman and then ignored a notice requiring the equipment to be used only in connection with other products made by the patentee, was not liable for infringement.[53] In particular, the Court expressed misgivings over patentee efforts to "extend the scope of its patent monopoly by restricting the use of it to materials necessary in its operation, but which are no part of the patented invention."[54] It was not until 1942, however, that the Court specifically crafted a "misuse" doctrine in *Morton Salt Co. v. G.S. Suppiger Co.*[55] The evidence in that case showed that Morton leased its patented machines on condition that lessees use the machines only in conjunction with (nonpatented) salt tablets made by the plaintiff's subsidiary.[56] Morton filed a patent infringement suit against another company that allegedly made and sold infringing machines.[57] The Court held Morton's patent unenforceable, despite evidence that (1) the defendant itself was not a party to the restrictive licenses, and was not charged with inducing others to breach the restrictive licenses, (2) the licenses may not have violated the Clayton Act, because they did not substantially lessen competition or tend to create a monopoly in the market for salt tablets, and (3) the defendant's machines may well have come within the scope of the

[52] 243 U.S. 502 (1917).

[53] *See id.* at 512-19.

[54] *See id.* at 516, 519.

[55] 314 U.S. 488 (1942).

[56] *See id.* at 489-91.

[57] *See id.* at 490-91.

patent's claims.[58] Citing the equitable principle that courts "may appropriately withhold their aid where the plaintiff is using the right asserted contrary to the public interest," the Court announced that the use of a patent "to suppress competition in the sale of an unpatented article may deprive the patentee of the aid of a court of equity to restrain an alleged infringement by one who is a competitor."[59] Significantly, the Court refused to limit the principle to cases in which the patentee's misconduct relates to the particular transaction at issue, or in which the defendant competes with the plaintiff in the market for the unpatented product.[60] Thus, "regardless of whether the particular defendant has suffered from the misuse of the patent," the patentee is foreclosed from obtaining relief, "at least until it is made to appear that the improper practice has been abandoned and that the consequences of the misuse . . . have been dissipated."[61]

Subsequent cases extended the holding of *Morton Salt* in various respects, holding for example that misuse could consist of setting royalty rates on the basis of the licensee's sale of unpatented products;[62] or licensing a patent on condition that the licensee continue paying royalties after the expiration of the patent.[63] More

---

[58] *See id.* at 490-92.

[59] *Id.* at 491, 492.

[60] *See Morton Salt*, 314 U.S. at 492-94. The doctrine of unclean hands, however, is much narrower. *See* RESTATEMENT (THIRD) OF RESTITUTION AND UNJUST ENRICHMENT § 32 cmt. d (Tent. Draft No. 3, 2004) (stating that, under the doctrine of unclean hands, "a party guilty of inequitable conduct *in the underlying transaction* may on that account be denied a claim based on unjust enrichment") (emphasis added).

[61] *Id.* at 493, 494; *see also* International Salt Co. v. United States, 332 U.S. 392 (1947); B.B. Chem. Co. v. Ellis, 314 U.S. 495 (1942).

[62] *See* Zenith Radio Corp. v. Hazeltine Res., Inc., 395 U.S. 100, 139 (1969); *but see* Engel Indus. v. Lockformer, 96 F.3d 1398, 1408-09 (Fed. Cir. 1996) (holding that "voluntary" package licensing does not constitute misuse, and distinguishing *Zenith* on the ground that the licensing arrangement was coerced).

[63] *See* Brulotte v. Thys Co., 379 U.S. 29, 30-34 (1965); Scheiber v. Dolby Labs., Inc., 293 F.3d 1014, 1017-21 (7th Cir. 2002); *cf.* Aronson v. Quick Point Pencil Co., 440 U.S. 257 (1979) (holding that federal patent

controversially, the Court in *Mercoid Corp. v. Mid-Continent Investment Co.* (*Mercoid I*)[64] held that patent misuse could arise from tying a patent license to the sale of a *non-staple* product--that is, a product that has no substantial use other than in connection with the patented invention[65]--thus creating a rule that tended to reduce the doctrine of contributory infringement to a nullity.[66] A 1952 amendment to the Patent Act,[67] however, as interpreted by the Supreme Court in 1980,[68] overruled *Mercoid I* and put to rest the notion that patentees cannot enforce their rights against persons who knowingly supply nonstaple articles for others to use for infringing purposes.[69] More recently, Congress enacted the Patent Misuse Reform Act of 1988[70] to further limit the application of the misuse doctrine as applied to tying arrangements. The act added two new subparts to § 271(d), which together require the defendant to prove that the patentee has market power, rather than to rely upon a presumption that the patent confers such power.[71]

Decisions of the United States Court of Appeals for the Federal Circuit, since 1982 the court that hears most appeals in patent cases, have added some clarity to the law of misuse while also leaving some abiding questions. The Federal Circuit describes

---

law did not preempt a contract to pay a 5% royalty indefinitely if no patent issued on the inventor's application, and a lower royalty for a limited time if the patent had issued).

[64] 320 U.S. 661 (1944). *See also* Mercoid Corp. v. Minneapolis Honeywell Regulator Co., 320 U.S. 680 (1944) (*Mercoid II*).

[65] *See Mercoid I*, 320 U.S. at 663-70.

[66] *See* Dawson Chem. Co. v. Rohm & Haas Corp., 448 U.S. 176, 199-200 (1984) (discussing *Mercoid I*).

[67] *See* Patent Act of 1952, ch. 950, 66 Stat. 792, 811 (codified at 35 U.S.C. § 271(d)(1)-(3)).

[68] *See Dawson Chem. Co.*, 448 U.S. 176.

[69] *See id.* at 200-23.

[70] *See* Pub. L. No. 100-703, § 201, 102 Stat. 4674, 4676 (codified at 35 U.S.C. §§ 271(d)(4), (5) (2000)).

[71] *See Independent Ink*, 396 F.3d at 1342 n.7.

misuse as an effort to "impermissibly broaden[ ] the 'physical or temporal scope' of the patent grant with anticompetitive effect,"[72] and it employs a three-step inquiry to determine whether a patentee is guilty of misuse. The first step is to characterize the practice at issue as per se misuse, per se lawful, or falling within some intermediate category.[73] Practices constituting per se misuse include tying—subject, however, to § 271(d)(5), which requires proof of market power and thus renders tying a per se offense in name only—and "arrangements in which a patentee effectively extends the term of its patent by requiring post-expiration royalties."[74] Per se lawful practices include those listed in § 271(d)(1)-(4).[75] If the practice is neither per se misuse nor per se lawful, then:

> a court must determine if that practice is "reasonably within the patent grant, *i.e.*, that it relates to subject matter within the scope of the patent claims." If so, the practice does not have the effect of broadening the scope of the patent claims and thus cannot constitute patent misuse. If, on the other hand, the practice has the effect of extending the patentee's statutory rights and does so with an anti-competitive effect, that practice must then be analyzed in accordance with the "rule of reason." Under the rule of reason, "the finder of fact must decide whether the questioned practice imposes an unreasonable restraint on competition, taking into account a variety of factors, including specific information about the relevant business, its condition before and after the restraint was imposed, and the restraint's history, nature, and effect."[76]

Using this approach, the Federal Circuit has held that practices such as threatening to void or limit warranties if third parties did not agree to purchase certain unpatented products from the patent owner, or requiring the licensee to acknowledge the validity of

---

[72] Windsurfing Int'l, Inc. v. AMF, Inc., 782 F.2d 995, 1001 (Fed. Cir. 1986) (quoting Blonder-Tongue Labs., Inc. v. University of Ill. Fdn., 402 U.S. 313, 343 (1971)).

[73] *See* Va. Panel Corp. v. MAC Panel Corp., 133 F.3d 860, 869 (Fed. Cir. 1998).

[74] *See id.*

[75] *See id.*; 35 U.S.C. §§ 271(d)(1)-(4).

[76] *Va. Panel*, 133 F.3d at 869 (citations omitted).

and avoid using the patentee's *trademarks*, did not constitute misuse because they did not broaden the scope of the patent claims.[77] Other practices, such as field-of-use restrictions, are evaluated, if at all, under the rule of reason.[78]

Perhaps the most controversial of the Federal Circuit's misuse decisions is *Mallinckrodt, Inc. v. Medipart, Inc.*[79] In *Mallinckrodt*, the patentee sold its patented medical devices to hospitals on condition that the latter would dispose of these devices after a single use.[80] Some hospitals ignored this restriction, however, sending their used devices to Medipart for reconditioning and then using them again.[81] The patentee filed suit against Medipart for infringement and for inducing infringement on the part of the hospitals.[82] Characterizing the condition as a field-of-use restriction, rather than as a per se offense such as tying or price-fixing, the court reversed and remanded for a determination whether the restraint was "reasonably within the patent grant, *i.e.*, that it relates to subject matter within the scope of the patent claims"--and only if the answer to that question was no, to determine whether the anticompetitive effects of the restraint violated the rule of reason.[83] The court also concluded that, if the restriction on reuse

---

[77] *Va. Panel*, 133 F.3d at 868-71 (warranties); *Windsurfing*, 782 F.2d at 1001-02 (trademarks).

[78] *See, e.g.*, U.S. Philips Corp. v. ITC, 424 F.3d 1179 (Fed. Cir. 2005) (holding that contracts requiring patent licensees to pay a fixed royalty rate for a package of patents, including some that the licensees deemed to be nonessential, was not per se misuse and passed muster under the rule of reason); B. Braun Med., Inc. v. Abbott Labs., 124 F.3d 1419, 1426-27 (Fed. Cir. 1997) (remanding to consider whether field-of-use restrictions broadened the scope of the patent, and if so whether they violated the rule of reason).

[79] 976 F.2d 700 (Fed. Cir. 1992).

[80] *See id.* at 701-02.

[81] *See id.* at 702.

[82] *See id.*

[83] *See id.* at 703-09.

proved to be lawful under the preceding analysis, then the first-sale doctrine also would provide no defense to the claim of unauthorized use.[84]

The extent to which the misuse doctrine departs from antitrust principles still remains unclear around the edges. On the one hand, both the Supreme Court and the Federal Circuit have continued to state that misuse and antitrust are not coextensive,[85] and at least in the context of per se misuse this appears to be correct. An agreement to continue collecting royalties after the patent term, for example, would not constitute an antitrust violation, absent proof of anticompetitive effect. And even though § 271(d)(5) moves the law of tying misuse closer to its antitrust counterpart, the overlap is not complete; it may be that assertions of tying misuse can be sustained on a lesser showing of anticompetitive harm than would be the case with respect to an analogous antitrust claim.[86] In addition, any patent defendant can raise the defense of misuse, even if he has suffered no harm from the conduct constituting misuse, whereas private antitrust claims

[84] *See id.* at 709. On this reasoning, the first-sale doctrine, which generally permits the lawful owner of a product incorporating a patented invention to use and resell the product without permission from the patent owner, would appear to be a default rule that the parties can modify at will. *See* Julie E. Cohen & Mark A. Lemley, *Patent Scope and the Software Industry*, 89 CAL. L. REV. 1, 33-36 (2001) (criticizing this result).

[85] *See* Zenith Radio Corp. v. Hazeltine Res., Inc., 395 U.S. 100, 140 (1969) (stating that, when there is patent misuse, "it does not necessarily follow that the misuse embodies the ingredients of a violation of either § 1 or § 2 of the Sherman Act"); C.R. Bard, Inc. v. M3 Sys., Inc., 157 F.3d 1340, 1372 (Fed. Cir. 1998) (stating that misuse is "a broader wrong than antitrust violation because of the economic power that may be derived from the patentee's right to exclude," and that it "may arise when the conditions of antitrust violation are not met"); *Mallinckrodt*, 976 F.2d at 704.

[86] *See, e.g.*, Senza-Gel Corp. v. Seiffhart, 803 F.2d 661, 670 (Fed. Cir. 1986) (holding that an antitrust tying claim involves two separate goods, as defined by consumer demand, being tied, whereas the law of misuse "need not look to consumer demand (which may be nonexistent) but need look only to the nature of the claimed invention as the basis for determine whether a product is a necessary concomitant of the invention or an entirely separate product"). Note, however, that in *C.R. Bard*, 157 F.3d at 1368, the court affirmed a jury verdict that a product design change that made it more difficult for purchasers of a product covered by a patent to obtain spare parts from anyone other than the patentee violated Sherman Act § 2, in light of evidence that the patentee had market power in the market for the replacement part, and that the design change reduced the product's technical functionality. *See id.* at 1381-83. The court reversed a finding that the patentee had committed misuse, however, on the grounds that (1) the jury instruction on misuse was vague, and (2) there was no evidence that the patentee's conduct constituted per se misuse or was "not 'reasonably within the patent grant.'" *Id.* at 1372-73.

can be asserted only by a party who himself has suffered, or is threatened with, "injury of the type the antitrust laws were meant to prevent."[87] Finally, although the Federal Circuit's adoption of a rule of reason taken from antitrust case law suggests that in non-per se cases the antitrust and misuse standards are identical, it remains to be seen whether the Supreme Court would agree, should it ever take up the issue.

That misuse may extend beyond the scope of antitrust is more apparent in the copyright context. Although the Supreme Court has never clearly endorsed a copyright misuse doctrine, beginning in the 1990s several federal courts of appeals have done so, beginning with the Fourth Circuit's decision in *Lasercomb America, Inc. v. Reynolds.*[88] In *Lasercomb*, the plaintiff owned a copyright in a computer-assisted die-making software program, Interact.[89] The defendant copied the program, and then created and marketed its own, nearly identical, competing program.[90] The district court entered judgment for the plaintiff on its copyright infringement claim, but the Fourth Circuit reversed on the ground of copyright misuse.[91] Citing the Supreme Court's patent misuse decisions as a basis for recognizing an analogous copyright misuse defense, the court took issue with provisions in Lasercomb's standard licensing agreement requiring licensees not to develop or market their own computer-assisted die-making software

[87] *See* Brunswick Corp. v. Pueblo Bowl-O-Mat, Inc., 429 U.S. 477, 489 (1977). There also is no necessary connection between the amount of harm suffered by the victims of the misuse and the penalty imposed upon the patentee, namely the loss of the right to enforce the patent. *See* Mark A. Lemley, *The Economic Irrationality of the Patent Misuse Doctrine*, 78 CAL. L. REV. 1599, 1616-17 (1990).

[88] 911 F.2d 970 (4th Cir. 1990).

[89] *See id.* at 971.

[90] *Id.*

[91] *See id.* at 979.

during the 99-year term of the agreement.[92] Although the defendant itself had never executed the standard agreement, the court concluded that the license was "anticompetitive," an "attempt[ ] to suppress any attempt by the licensee to independently implement the idea which Interact expresses" and "to use its copyright . . . to control competition in an area outside the copyright, i.e., the idea of computer-assisted die manufacture . . . ."[93] The court furthermore rejected recourse to antitrust's rule of reason, stating that "the question is not whether the copyright is being used in a manner violative of antitrust law (such as whether the licensing agreement is 'reasonable'), but whether the copyright is being used in a manner violative of the public policy embodied in the grant of a copyright."[94]

Other courts followed suit in the wake of *Lasercomb*. For example, the Fifth Circuit in *Alcatel USA, Inc. v. DGI Technologies, Inc.*[95] held that an agreement requiring licensees of the plaintiff's software to use the software only in conjunction with hardware manufactured by the plaintiff constituted misuse[96] (while at the same time holding that the evidence that the plaintiff had power over a properly defined market was insufficient

[92] *See id.* at 973, 978-79.

[93] *Id.* at 973, 977, 978, 979.

[94] *Id.* at 977, 978.

[95] 166 F.3d 772 (5th Cir. 1999).

[96] *See id.* at 777, 793-94. Although the court described the licensing provision noted above as the basis for misuse, it also stated that the defendant, in order to ensure that the hardware it was developing would be compatible with the plaintiff's operating systems software, had to reverse engineer that software. *See id.* 793-94 (stating that "If [Alcatel] is allowed to prevent such copying, then it can prevent anyone from developing a competing microprocessor card, even though it has not patented the card"). This reasoning suggests that the fair use analysis developed in cases such as *Sega Enterprises, Ltd. v. Accolade, Inc.*, 977 F.2d 1510 (9th Cir. 1992), might have been dispositive, but the *Alcatel* court did not discuss fair use.

to support the defendant's counterclaim for violation of § 2 of the Sherman Act).[97] Similarly, the Ninth Circuit held in *Practice Management Information Corp. v. American Medical Association*[98] that the AMA committed misuse by requiring a federal agency to "use, copy, publish, and distribute" the AMA's copyrighted medical procedure codes for use in administering Medicare claim forms, on condition that the agency agree "not to use any other system of procedure nomenclature . . . for reporting physicians' services."[99] The court found the "adverse effects" of this requirement to be "apparent," because it "gave the AMA a substantial and unfair advantage over competitors"[100]--even though there was "no evidence that anyone wishing to use the CPT has any difficulty in obtaining it,"[101] or that the AMA's copyright posed any "realistic threat to public access"[102] or "stifle[d] independent creative expression in the medical coding industry,"[103] or that the AMA had violated the antitrust laws.[104]

Another interesting variation is presented in *Assessment Technologies of Wisconsin, LLC v. WIREdata, Inc.*[105] At issue was an attempt on the part of WIREdata, a company that serviced real estate brokers, to obtain information concerning properties

---

[97] *See id.* at 783-84.

[98] 121 F.3d 516 (9th Cir. 1997).

[99] *See id.* at 517.

[100] *Id.* at 521.

[101] *Id.* at 519.

[102] *Id.*

[103] *Id.* at 520 n.8. On this basis, the court concluded that the merger doctrine did not apply. *See id.*; *see also infra* Part III.B (discussing merger).

[104] *See id.* at 521.

[105] 350 F.3d 640 (7th Cir. 2003).

located in several Wisconsin towns.[106] The towns themselves collected this information for purposes of property-tax assessments from independent contractor assessors who coded the information using a program developed by Assessment Technologies (AT).[107] Each time an assessor used the program, the program created a copyrightable compilation of data organized into 34 master categories and 456 subcategories; but the raw data themselves consisted only of uncopyrightable facts.[108] When WIREdata sought the data from the municipalities under Wisconsin's open-records law, AT filed suit to enjoin WIREdata from inducing the towns to infringe AT's copyright.[109] The district court granted the injunction but the Seventh Circuit, in an opinion by Judge Posner, reversed.[110] Although the principal ground for reversal was that AT's copyright extended only to the original aspects of its compilation, and not the raw data, and therefore there could be no infringement for merely supplying those data,[111] the court made two further observations. First, the court cited *Sega Enterprises Ltd. v. Accolade, Inc.*[112] for the proposition that fair use would shield the municipalities if it were impossible for them to extract the data

[106] *See id.* at 642.

[107] *See id.* at 642-44.

[108] *See id.* at 642-43.

[109] *See id.* at 642.

[110] *See id.* at 642, 648.

[111] *See id.* at 643-44. Copyright subsists in original works of authorship, fixed in a tangible medium of expression. *See* 17 U.S.C. § 102(a). Copyright does not subsist in facts, but it may subsist in a compilation that selects or arranges facts in a manner that demonstrates minimal creativity. *See* 17 U.S.C. §§ 101, 102(b); Feist v. Rural Tel. Servs., 499 U.S. 340, 348-49 (1991). Thus, one who copies individual facts from a copyrighted compilation, but not the original selection or arrangement, does not infringe the compilation copyright. *Feist*, 499 U.S. at 348. In *Assessment Technologies*, the selection of facts appears to have been comprehensive, and therefore not original, but AT's arrangement in accordance with its 456 categories probably demonstrated sufficient creativity to satisfy originality. All that WIREdata wanted to copy, however, were the facts themselves. *See Assessment Technologies*, 350 F.3d at 644.

[112] 977 F.2d 1510 (9th Cir. 1992). The case is discussed herein at notes __ and accompanying text.

without copying the copyrightable arrangement.[113] Second, the court asserted, any attempt by AT to use its software license, which arguably forbade the towns from disseminating any data collected with the use of the program, so as to prevent them "from revealing *their own* data . . . might constitute misuse."[114] The court noted that there was no evidence that AT possessed market power, but it concluded that this did not matter:

> The argument for applying copyright misuse beyond the bounds of antitrust . . . is that for a copyright owner to use an infringement suit to obtain property protection, here in data, that copyright law clearly does not confer, hoping to force a settlement or achieve an outright victory over an opponent that may lack the resources or the legal sophistication to resist effectively, is an abuse of process.[115]

As the above discussion demonstrates, there are two overarching policies that form the contours of the misuse doctrine. One is competition policy; but the case law is largely silent as to why competition policy demands that antitrust be supplemented in this fashion, or why society should condemn, on competition grounds, conduct that falls short of constituting an antitrust violation. The other is what I referred to in the Introduction as the beyond-the-scope rationale: the concern that IP owners should not be allowed to "extend the scope" of their monopoly. This concern is clearly related to competition policy, insofar as the scope of IP rights can be understood as reflecting some balance of monopoly incentives and competition. Defenders of the misuse doctrine, however, might argue that that balance is sometimes different from the balance struck by antitrust. IP policy dictates that facts and ideas, for example, must remain in the public domain,

---

[113] *See id.* at 644-45.

[114] *See id.* at 642, 644, 646.

[115] *Id.* at 647. This statement contrasts with Judge Posner's more skeptical thoughts on misuse as expressed in *USM Corp. v. SPS Technologies, Inc.*, 694 F.2d 505, 512 (7th Cir. 1982) (stating that "[o]ur law is not rich in alternative concepts of monopolistic abuse; and it is rather late in the day to try to develop one without in the process subjecting the rights of patent holders to debilitating uncertainty").

regardless of whether the exclusive control over some facts or ideas would permit the exercise of market power in a given case.

The preceding analysis is useful, but only up to a point. An initial difficulty is that the beyond-the-scope rationale offers no reason why efforts to expand one's IP rights beyond the scope of the grant is such an egregious offense that, contrary to conventional standing requirements, it may be pleaded by someone who is unaffected by the offending transaction.[116] In addition, the beyond-the-scope rationale often appears to rest upon the premise that the scope of the grant is clear, but this is often not the case. Although the claims of a patent are described as the "metes and bounds" of the invention,[117] IP rights don't come with metes and bounds in any literal sense; determining whether an invention falls within the literal scope of those claims is often a difficult task.[118] Determining the scope of a copyright is even more difficult, because there is nothing analogous to patent claims that define the outer limits of the work of authorship. To be sure, that outer limit must stop short of mere ideas and facts. In this respect, *Assessment Technologies* was an easy case, because the data at issue were facts. As the following subsection shows, however, it is not always easy to determine what an idea or fact is; what falls within the category of idea or fact is itself often *not* a fact, but rather a conclusion based on policy.[119] All the more so are such determinations as whether the scope of the grant prevents the IP owner from requiring licensees to pay post-expiration royalties, or to

[116] *See* Lemley, *supra* note __, at 1614-20; *cf.* Dan L. Burk, *Anticircumvention Misuse*, 50 UCLA L. REV. 1095, 1123 (2003) (arguing that misuse doctrine helps to preserve the integrity of the IP system, but not articulating why this policy requires parties who would not have standing to assert the unclean hands defense to have standing to assert misuse).

[117] *See* Univ. of Rochester v. G.D. Searle & Co., 375 F.3d 1303, 1308 (Fed. Cir. 2004).

[118] *See* Brooks Furniture Mfg., Inc. v. Dutailer Int'l, Inc., 393 F.3d 1378, 1384 (Fed. Cir. 2005).

[119] *See infra* notes __ and accompanying text.

purchase other goods and services, or to waive certain defenses. To conclude that permitting such conduct would enable IP owners to expand the scope of their "monopoly" assumes that the monopoly does not encompass the right to engage in such conduct in the first place. Whether that conclusion is correct should be decided on the basis of consequences, not formal distinctions.[120]

B. Merger and Fair Use

The merger doctrine arises from the distinction that copyright law makes between (uncopyrightable) ideas and (copyrightable) expression. An early case often cited as having recognized this distinction, *Baker v. Selden*, involved a dispute between the authors of two books explaining bookkeeping systems.[121] The Selden book explained one of the systems and included forms consisting of ruled lines and headings to illustrate that system; Baker's book explained a similar system and used somewhat similar, though hardly identical, forms.[122] In ruling for the defendant, the Supreme Court reasoned, first,

---

[120] Cases such as *Brulotte v. Thys Co.*, 379 U.S. 29 (1965), for example, which condemned post-expiration royalties, are often (rightly) criticized for elevating form over substance. Nonparties to the contract at issue in *Brulotte* remained free to practice the patent at issue following its expiration. Moreover, the principal consequence of invalidating post-expiration royalty contracts is simply to preclude the parties from agreeing to an efficient financing method similar to an installment contract. *See* LANDES & POSNER, *supra* note __, at 380-81, 417-18 (describing *Brulotte* as "one of the all-time economically dumb Supreme Court decisions"); Ian Ayres & Paul Klemperer, *Limiting Patentees' Market Power Without Reducing Innovation Incentives: The Perverse Benefits of Uncertainty and Non-Injunctive Remedies*, 97 MICH. L. REV. 985, 1026-27 (1999).

[121] 101 U.S. 99 (1879). *Baker v. Selden* itself never uses the term "expression," however, and Pamela Samuelson argues that the case is best viewed as resting upon the distinction between patentable and copyrightable subject matter. *See* Pamela Samuelson, *The Story of* Baker v. Selden, *in* INTELLECTUAL PROPERTY STORIES __, __ (Jane C. Ginsburg & Rochelle Cooper Dreyfuss eds. 2005).

[122] *See Baker*, 101 U.S. at 99-101. Scholars have argued that the case could have been decided on the narrower ground that Baker's forms were *not* substantially similar to Selden's. *See* 1 MELVILLE B. NIMMER & DAVID NIMMER, NIMMER ON COPYRIGHT § 2.18[B][1], at 2-202 (2005); Samuelson, *supra* note __, at __. The Court's expansive ruling responded to an argument that copyright law afforded Selden the exclusive right to the use of his system. *See* 1 NIMMER & NIMMER, *supra*, § 2.18[B][2], at 2-202; Samuelson, *supra* note __, at __.

that a bookkeeping system as such—as opposed to the words that Selden used to explain his system—did not fall within the subject matter of copyright.[123] Second, the Court asserted that Selden's forms fell outside the scope of copyright, to the extent those or similar forms "must necessarily be used as incident to" the system.[124]

The meaning of *Baker v. Selden* has long been the subject of varying opinions. The most persuasive reading of the case, however, is that one is privileged to copy expression that is a "necessary incident" to the use of an uncopyrightable idea[125]--or, to put it another way, that a copyright owner cannot leverage her copyright over expression into control over uncopyrightable subject matter (ideas).[126] This last interpretation, in turn, elides into the merger doctrine, the first articulation of which arose in *Morrissey v. Procter & Gamble Co.*[127] The plaintiff in *Morrissey* owned the copyright to a set of rules for a sweepstakes game involving contestants' Social Security numbers; the defendant copied the rules for use in its own contest.[128] In finding for the defendant, the First Circuit reasoned that:

> [w]hen the uncopyrightable subject matter is very narrow, so that 'the topic necessarily requires', if not only one form of expression, at best only a limited number, to permit copyrighting would mean that a party or parties, by copyrighting a mere handful of forms, could exhaust all possibilities of future use of the substance.[129]

---

[123] *See Baker*, 101 U.S. at 101-04. This outcome is preserved in § 102(b) of the Copyright Act, which states that copyright does not subsist in, among other things, systems. *See* 17 U.S.C. § 102(b).

[124] *Baker*, 101 U.S. at 104.

[125] *See* Samuelson, *supra* note __, at __.

[126] *See* Bucklew v. Hawkins, Ash, Baptie & Co., LLP, 329 F.3d 923, 928 (7th Cir. 2003); 4 NIMMER & NIMMER, *supra* note __, § 13.03[B][3], at 13-78.

[127] 379 F.2d 675 (1st Cir. 1967).

[128] *See id.* at 676, 678.

[129] *Id.* at 678.

Other courts have applied the merger doctrine to permit the unauthorized copying of items including a piece of jewelry depicting a bee,[130] a map based upon a public domain source,[131] a municipal building code,[132] and elements of computer programs that are dictated by functional considerations or other external constraints.[133] In a variety of other cases, however--some involving similar subjects, and others involving items such as data compilations,[134] estimates of current or future prices,[135] and religious scriptures[136]--courts have held that the relevant ideas did not merge with the plaintiff's expression.

Some applications of the merger doctrine seem fairly intuitive. For example, in *Kern River Gas Transmission Co. v. Coastal Corp.*, the Fifth Circuit held that the doctrine applied to maps made from government surveys, over which the plaintiff had superimposed lines portraying its proposed route for a natural gas pipeline.[137] According to the plaintiff's own engineer, there was no other way of expressing the idea of the

---

[130] *See* Herbert Rosenthal Jewelry Corp. v. Kalpakian, 446 F.2d 738, 742 (9th Cir. 1971).

[131] *See* Kern River Gas Transmission Co. v. Coastal Corp., 899 F.2d 1458, 1463-64 (5th Cir. 1990); *but see* Mason v. Montgomery Data, Inc., 967 F.2d 135, 138-41 (5th Cir. 1992) (holding that merger did not apply to the maps in that case).

[132] *See* Veeck v. S. Bldg. Code Congress Int'l, Inc., 293 F.2d 791, 800-02 (5th Cir. 2002) (en banc).

[133] *See, e.g.*, Lexmark Int'l, Inc. v. Static Control Components, Inc., 387 F.3d 522, 537-42 (6th Cir. 2004).

[134] *See, e.g.*, Kregos v. Associated Press, 937 F.2d 700, 705-07 (2d Cir. 1991); American Dental Ass'n v. Delta Dental Plans Ass'n, 126 F.3d 977, 979 (7th Cir. 1997).

[135] *See, e.g.*, CDN, Inc. v. Kapes, 197 F.3d 1256, 1262 (9th Cir. 1999); CCC Info. Servs. v. Maclean Hunter Mkt. Reports, 44 F.3d 61, 68-73 (2d Cir. 1994).

[136] *See* Merkos L'Inyonei Chinuch, Inc. v. Otsar Sifrei Lubavitch, Inc., 312 F.3d 94, 99 (2d Cir. 2002); Penguin Books U.S.A., Inc., v. New Christian Church of Full Endeavor, 55 U.S.P.Q. 2d (BNA) 1680, 1696 (S.D.N.Y. 2000); Religious Tech. Center v. Lerma, 40 U.S.P.Q.2d (BNA) 1569, 1573 (E.D. Va. 1996).

[137] *See Kern River*, 899 F.2d at 1464.

proposed route.[138] But other applications of merger are less obvious. For one thing, the premise that only a "limited number" of ways exist to express a given idea rests upon the assumptions that (1) there is some discernible meaning to the term "limited number," and (2) courts can determine whether small differences are sufficiently distinct to constitute separate ways of expressing an idea. But there is little guidance on these issues from the case law, and critics complain that courts sometimes leap to the conclusion that merger applies, on the basis of insubstantial evidence.[139]

A second, more fundamental, problem is that there is no clear formula for deciding, as a first step before assessing the number of ways to express an idea, precisely what the relevant idea *is*. If the idea is defined broadly enough--say, the idea for a game of chance, as opposed to the idea for a sweepstakes game involving Social Security numbers—there will always be a multiplicity of ways of expressing that idea. Deciding how broadly or how narrowly to define the idea determines the outcome, but courts rarely disclose what factors lead them to seize upon one definition of "idea" over another.[140] A few courts and commentators, however, have recognized that the task of defining the idea must derive from policy considerations, including whether the interest in accessing the plaintiff's work is so great that competitive need outweighs copyright's incentive

---

[138] *See id.* at 1464.

[139] Several scholars, for example, have taken issue with *Morrissey*'s conclusion that there only a handful of ways to draft sweepstakes instructions. *See* 1 NIMMER & NIMMER, *supra* note __, § 2.18[C][2], at 2-204.5; Michael Abramowicz, *Copyright Redundancy*, *available at* http://ssrn.com/abstract_id=374580, at 25.

[140] The Second Circuit has stated that it will construe the "idea" more broadly when it reflects "taste or personal opinion," *Kregos*, 937 F.2d at 707, and more narrowly when it comprises the "building blocks of understanding," *CCC*, 44 F.3d at 71. In theory, this standard correctly leaves the "building blocks" less vulnerable to capture by private parties; but it still leaves a good deal of play in the joints. *See, e.g.*, 4 NIMMER & NIMMER, *supra* note __, § 13.03[B][3], at 13-78.4 to -78.5 (expressing surprise at the court's application of its own test in *CCC*).

scheme.[141] But the process of idea definition remains primitive in comparison with the process of market definition in antitrust law, even though both arise from a common interest in promoting competition. As with the misuse doctrine, courts rarely explain why a specific outcome is necessary to achieve a desirable competitive balance.[142]

Similar, in many respects, are a subset of cases decided under the fair use doctrine. As a general matter the fair use doctrine, as embodied in § 107 of the Copyright Act, entitles one to engage in otherwise infringing activity subject to certain qualifications. Section 107 begins with a list of uses that, all other things being equal, are more likely to be deemed fair, including uses for purposes of "criticism, comment, news reporting, teaching . . . scholarship, or research."[143] But not all uses that fall within these categories are necessarily fair, while some uses that fall without are.[144] The statute goes on to enumerate four factors that are relevant to the fair use determination, including the nature and purpose of the use; the nature of the work; the amount and substantiality of the

[141] A few decisions at least refer to competitive need as a factor in their analysis. *See Kalpakian* ("The guiding consideration in drawing the line is the preservation of the balance between competition and protection reflected in the patent and copyright laws"); *CCC*, 44 F.3d at 72 n.25 (citing *Kalpakian*); *Kern River*, 899 F.2d at 1463; Apple Computer, Inc. v. Franklin Computer Corp., 714 F.2d 1240, 1253 (3d Cir. 1983). As for commentary, *see, e.g.*, 4 NIMMER & NIMMER, *supra* note __, § 13.03[f][1], at 13-124 (stating that the line between ideas and expression "is a pragmatic one, drawn not on the basis of some metaphysical property of 'ideas,' but by balancing the need to protect the labors of authors with the desire to assure free access to ideas"); Jane C. Ginsburg, *No "Sweat"? Copyright and Other Protection of Works of Information after* Feist v. Rural Telephone, 92 COLUM. L. REV. 338, 346 (1992) ("In copyright law, an 'idea' is not an epistemological concept, but a legal conclusion prompted by notions—often unarticulated and unproven—of appropriate competition"); John Shepard Wiley, Jr., *Copyright at the School of Patent*, 58 U. CHI. L. REV. 119, 158-59 (1991) (proposing that courts apply merger when doing so would not risk deterring authors from creating); *see also* Thomas F. Cotter, *Gutenberg's Legacy: Copyright, Censorship, and Religious Pluralism*, 91 CAL. L. REV. 323, 362-63 (2002); Shubha Ghosh, *Legal Code and the Need for a Broader Functionality Doctrine in Copyright*, 50 J. COPYRIGHT SOC'Y U.S.A. 71, 101-10 (2003).

[142] *See* Ghosh, *supra* note __, at 102 (making a similar point).

[143] 17 U.S.C. § 107 (2000).

[144] *Compare* Sony Corp. v. Universal City Studios, 464 U.S. 417, 442 (1984) (holding that private copying for television programs in order to watch them at more convenient times constituted fair use per se), *with* Princeton Univ. Press. v. Michigan Doc. Servs. 99 F.3d 1381, 1383 (6th Cir. 1996) (en banc) (holding that copying for purposes of making college course packets was not fair use).

portion used, in relation to the copyrighted work as a whole; and the effect of the use upon the market for or value of the copyright plaintiff's work.[145]

Some applications of fair use have relatively little to do with competition policy as such. One class of fair uses, for example, involves cases in which the copyright owner probably would have granted the defendant permission to copy, but the cost of negotiating for permission would have exceeded the value of that use and therefore would have precluded voluntary bargaining.[146] In such a case, excusing the defendant from liability makes users better off, and owners no worse off, than they otherwise would have been. Another class involves uses that are perceived to confer social benefits substantial enough to warrant departure from the private bargaining model. The use of quotations from a work in order to critique it or to report upon it, for example, presumably gives rise to social benefits from the resulting commentary; at the same time, exempting the user from liability prevents copyright owners from using their copyrights to deflect criticism.[147] A subset of this latter class comprises cases in which courts invoke competitive need as a reason for permitting copying. Two of the leading cases,

[145] *See* 17 U.S.C. 107.

[146] *See* LANDES & POSNER, *supra* note __, at 115-16; Thomas F. Cotter, *Accommodating Religious Uses of Copyrighted Works Under the Fair Use Doctrine and Copyright Act § 110(3)*, 22 CARDOZO ARTS & ENTER. L.J. 43, 44-45 (2004).

[147] *See* LANDES & POSNER, *supra* note __, at 117-23; Cotter, *supra* note __, at 45-47. Conceived more broadly, fair use may confer distributive benefits--for example, by exempting users from having to pay for content in situations in which ability to pay may constrain willingness to pay—or may prevent content owners from exerting control over new copying technologies. *See, e.g.*, Molly Schaffer Van Houweling, *Distributive Values in Copyright*, 83 TEX. L. REV. 1535, 1543-45 (2005).

Sega Enterprises Ltd. v. Accolade, Inc.[148] and *Sony Computer Entertainment, Inc. v. Connectix Corp.*,[149] involve the reverse engineering of computer software.

In *Sega*, the plaintiff Sega Enterprises marketed the Genesis video game console and an array of complementary video game cartridges.[150] It also licensed to video game developers the computer code that controlled the Genesis system, to enable the developers to create Genesis-compatible games; in return, the developers agreed that Sega would have the exclusive right to manufacture and distribute the games in exchange for periodic royalties.[151] Unwilling to agree to these terms, Accolade lawfully acquired copies of some Sega games; copied the games' object code, which included some copyrightable expression; and reverse-engineered that code to discover the uncopyrightable interface specifications that enabled compatibility with the Genesis console.[152] Using this information, Accolade created its own Genesis-compatible games, which did not incorporate Sega's copyrighted expression.[153] Sega nevertheless filed suit, claiming that the intermediate copying, even though done for purposes of creating the

[148] 977 F.2d 1510 (9th Cir. 1993).

[149] 203 F.3d 596 (9th Cir. 2000). *See also* Atari Games Corp. v. Nintendo of Am., 975 F.2d 832, 842-44 (Fed. Cir. 1992) (stating that copying for the purpose of reverse engineering lawfully-obtained code, so as to extract the uncopyrightable elements from it, is fair use). Similar to these reverse-engineering cases is *Lotus Dev. Corp. v. Borland Int'l, Inc.*, 49 F.3d 807 (1st Cir. 1995), *aff'd by an equally divided Court*, 516 U.S. 233 (1996), in which the First Circuit concluded that Borland's copying of Lotus's user interface, for the purpose of facilitating consumers' use of Borland's own spreadsheet program, did not infringe because the interface was an uncopyrightable method of operation. *See id.* at 815-19; *see also id.* at 821-22 (Boudin, J., concurring) (arguing that, in the alternative, Borland's copying may have been a fair use, because Borland copied for the purpose of marketing its independently-created product).

[150] *See* 977 F.2d at 1514.

[151] *See id.*

[152] *See id.* at 1514-16.

[153] *See id.* at 1516.

non-infringing, compatible games, itself constituted copyright infringement.[154] The Ninth Circuit disagreed, holding that Accolade's intermediate copying was a fair use.[155]

Of particular importance was the court's analysis of the fourth fair use factor, the "effect of the use on the market for or value of the copyrighted work"[156]--here, the Sega video games that Accolade copied. The court concluded that Accolade's conduct did not threaten to "diminish[ ] potential sales, interfer[e] with marketability, or usurp the market" by "supplanting" Sega games, but instead "simply enable[d] the copier to enter the market for works of the same type as the copied work."[157] Specifically:

> Accolade did not attempt to "scoop" Sega's release of any particular game or games, but sought only to become a legitimate competitor in the *field* of Genesis-compatible video games. Within that *market*, it is the characteristics of the game program as experienced by the user that determine the program's commercial success. . . .
>
> By facilitating the entry of a new competitor, the first lawful one that is not a Sega licensee, Accolade's disassembly of Sega's software undoubtedly "affected" the *market* for Genesis-compatible games in an indirect fashion. We note that . . . video game users typically purchase more than one game. There is no basis for assuming that Accolade's "Ishido" has significantly affected the *market* for Sega's "Altered Beast", since a consumer might easily purchase both; nor does it seem unlikely that a consumer particularly interested in sports might purchase both Accolade's "Mike Ditka Power Football" and Sega's "Joe Montana Football", particularly if the games are, as Accolade contends, not substantially similar. In any event, an attempt to monopolize the *market* by making it impossible for others to compete runs counter to the statutory purpose of promoting creative expression and cannot constitute a strong equitable basis for resisting the invocation of the fair use doctrine. . . .[158]

---

[154] *See id.*

[155] *See id.* at 1527-28.

[156] 17 U.S.C. § 107.

[157] 977 F.2d at 1523.

[158] *Id.* at 1523-24 (emphasis added).

Notice how the court, in the preceding paragraphs, initially asserts (twice) that the relevant "market" for fair use purposes is the market for "Genesis-compatible video games" (plural), only to follow up with a reference to the "market" for a *single* Sega game, "Altered Beast." In between, the court asserts that consumers "might easily purchase" both Sega and Accolade games, though without any analysis of consumers' budgetary constraints, or any citation to evidence that consumers would not view the parties' games as adequate substitutes for one another (which they clearly *could* be, even if the games were not substantially similar for purposes of infringement analysis). None of this necessarily means that the court reached the wrong result, but its conclusory market definition and its assertions about consumer demand are striking in comparison with the more thoughtful analysis these issues typically would receive in an antitrust suit.

In the second case, *Sony*, the defendant copied Sony's game console to develop a *platform*, the Virtual Game Station, to compete against Sony's PlayStation platform--rather than, as in *Sega*, to develop applications that would be complementary to a platform.[159] Again, the Ninth Circuit concluded that the use was fair, stating in relation to the fourth factor:

> . . . the Virtual Game Station is a legitimate competitor in the *market* for platforms on which Sony and Sony-licensed games can be played. . . . Sony understandably seeks control over the *market* for devices that play games Sony produces or licenses. The copyright law, however, does not confer such a monopoly.[160]

As in *Sega*, the court did not provide support for its definition of the relevant market--the "market for platforms on which Sony and Sony-licensed games can be played"—or for

[159] *See Sony*, 203 F.3d at 598. Specifically, the PlayStation console enabled users to play games on their television sets; Connectix's Virtual Game Station enabled them to play Sony-compatible games on their personal computers, though with some degradation of the visual aspects of the games.

[160] *Id.* at 607 (emphasis added).

its implicit assertion that Sony possessed power over a well-defined market. In fact, other video-game platforms compete vigorously against Sony's PlayStation.[161]

Once again, the preceding criticism does not necessarily mean that either case was wrongly decided. Both courts were careful to note that the fair use privilege applies only if the defendant has a "legitimate reason" for copying, and where copying is "necessary" to understand how a product works.[162] Moreover, both results arguably can be defended in light of the *Baker v. Selden* principle that a copyright owner cannot leverage its copyright so as to create a bottleneck over ideas. In fairness, in both of the Ninth Circuit cases at issue, the court *did* cite this principle as well in connection with its discussion of the first fair use factor;[163] and perhaps the principle that copying for the purpose of gaining access to the uncopyrightable elements of a work is sufficient to achieve the result in these cases, under the beyond-the-scope rationale, without consideration of market power.[164] Stated in such an absolute way, however, the principle is not without critics. Companies often invest much more in the creation of their platforms than in the creation of the applications for those platforms.[165] From an economic standpoint, the defendant's marketing of competing applications in *Sega* may have benefited the plaintiff, even if contrary to the court's assumption this marketing displaced some sales

---

[161] *See* Matt Richtel, *A Game Console for the Rest of Us*, N.Y. TIMES, Oct. 11, 2005, at C1 (discussing competition among Sony, Microsoft, and Nintendo).

[162] *See Sony*, 203 F.3d at 603-07; *Sega*, 977 F.2d 1526-28; *see also Atari*, 975 F.2d at 843.

[163] *See Sony*, 203 F.3d at 603-04; *Sega*, 977 F.2d at 1526-28.

[164] *See Sega*, 977 F.2d at 1527-28 ("where disassembly is the only way to gain access to the ideas and functional elements embodied in a copyrighted computer program and where there is a legitimate reason for seeking such access, disassembly is a fair use . . . as a matter of law").

[165] *See, e.g.*, Philip J. Weiser, *The Internet, Innovation, and Intellectual Property Policy*, 103 COLUM. L. REV. 534, 562 (2003) (stating that Sony invested $600 million to develop the Play Station, while Connectix spent $150,000 to develop its emulator).

of Sega's own applications, by encouraging game enthusiasts to use the Sega platform.[166] Marketing a competing platform in *Sony*, by contrast, threatened to undermine Sony's investment in the development of its PlayStation.[167] Critics such as Philip Weiser therefore argue that fair use should exempt intermediate copying in cases such as *Sega* but not in *Sony*.[168] Moreover, the beyond-the-scope rationale still leaves unresolved the issues of how to define the relevant idea or fact in the first place; perhaps some interface specifications are uncontroversially classifiable as "ideas," but other cases may not lend themselves to such ready conclusions.[169] Finally, as with misuse, there is a degree of question-begging that beyond-the-scope ignores. Why shouldn't the IP owner be permitted to exercise its rights, even when this makes it difficult for others to access ideas and facts? If the answer is to promote competition, how do we know that competition requires this result, absent an analysis of the relevant market?

C. Functionality

The functionality doctrine in trademark law presents yet another example of a standard that appears to be rooted in a competition policy more nebulous than anything found within the law of antitrust. Although functionality questions arise only in a subset

---

[166] *See* Pamela Samuelson & Suzanne Scotchmer, *The Law and Economics of Reverse Engineering*, 111 YALE L.J. 1575, 1616-18, 1622 (2002).

[167] *See* Maureen A. O'Rourke, *Toward a Doctrine of Fair Use in Patent Law*, 100 COLUM. L. REV. 1177, 1224 (2000). Conceivably, though, it also might have increased sales of Sony's applications, and the evidence showed that Sony was recouping its investment in the platform from sales of applications and not from sales of the platform, which was selling below cost. *See* Samuelson & Scotchmer, *supra* note ___, at 1622 n.220; Weiser, *supra* note __, at 603 n.295.

[168] *See* Weiser, *supra* note ___, at 602-03.

[169] Even deciding what a "fact" is can create difficulties, as the cases involves estimates of used car or coin prices show. *See CDN*, 197 F.3d at 1262; *CCC*, 44 F.3d at 67.

of trademark cases, principally those involving product design trade dress, these cases account for a growing percentage of trademark registrations and litigation.

Under U.S. law, a trademark is any distinctive symbol that identifies a unique product or service.[170] The principal right that flows from the ownership of a mark is to prevent another from the commercial use of the same or a similar mark, such that the other's use would be likely to confuse a substantial portion of prospective buyers into believing that the products have a common source or sponsor.[171] According exclusive rights over source-identifying symbols tends to reduce consumers' costs of distinguishing among products that appear similar but have different intrinsic qualities; as a corollary, trademark protection provides an incentive for producers to maintain a consistent level of quality in their products, so that the mark conveys a meaningful signal to consumers.[172]

Although some of the best-known trademarks are words (e.g, COCA-COLA), virtually any subject matter can serve as a source-identifying symbol, including colors,[173] sounds,[174] and "trade dress"—a term that can refer to product packaging, design, and perhaps other attributes as well.[175] Courts nevertheless have required persons asserting trademark rights in such nontraditional subject matter to satisfy two conditions which do not (usually) apply to more common fare, such as word marks. The first condition, imposed initially at common law, was that trade dress is protectable only if the putative

[170] *See* 15 U.S.C. § 1127 (2000); RESTATEMENT (THIRD) OF UNFAIR COMPETITION § 9 (1995).

[171] *See* 15 U.S.C. §§ 1114(1), 1125(a)(1)(A); RESTATEMENT (THIRD) OF UNFAIR COMPETITION § 20 (1995).

[172] *See* RESTATEMENT (THIRD) OF UNFAIR COMPETITION § 9 cmt. c (1995).

[173] *See* Qualitex Co. v. Jacobson Prods. Co., 514 U.S. 159, 162-66 (1995).

[174] *See* U.S. Trademark No. 1,395,550 (issued June 3, 1986) (registering "a lion roaring" to MGM Corp. for use in motion picture films and video tapes).

[175] *See* Wal-Mart Stores, Inc. v. Samara Bros., 529 U.S. 205, 209 (2000).

owner can demonstrate "secondary meaning," that is, that a substantial portion of consumers have come to associate the trade dress with a unique product.[176] In 1992, the U.S. Supreme Court ostensibly eliminated this requirement, as a matter of federal trademark law, as long as the trade dress at issue is "inherently distinctive."[177] Nine years later, however, the Court held in *Wal-Mart Stores, Inc. v. Samara Bros.* [178] that product *design* trade dress, as opposed to product *packaging* trade dress, never *is* inherently distinctive and therefore always must be supported by proof of secondary meaning.[179] A second condition is that only "nonfunctional" symbols can serve as trademarks.[180] In *Qualitex Corp. v. Jacobson Products Co.*, [181] a 1995 decision in which the Supreme Court held that distinctive, nonfunctional color can serve as a trademark, the Court stated that a product feature is functional if it is "'essential to the use or purpose of the article, or if it affects the cost or the quality of the article,' that is, if exclusive use of the feature would put competitors at a significant non-reputation-related disadvantage."[182] As a result, even if consumers associate a functional product attribute as emanating from a single firm, the attribute cannot be the subject of exclusive trademark rights; at best, confusion will have to be dispelled by some means, such as a disclaimer, short of

[176] *See* RESTATEMENT (THIRD) OF UNFAIR COMPETITION, §§ 13(b), 14 (1995) (defining secondary meaning); *id.* § 16 cmt. a (discussing the common-law rule). By contrast, a firm may assert trademark rights in a symbol that is "inherently distinctive" without proof of secondary meaning. *See id.* §§ 13(a), 18.

[177] *See* Two Pesos, Inc. v. Taco Cabana, Inc., 505 U.S. 763, 776 (1992).

[178] 529 U.S. 205 (2000).

[179] *See id.* at 216.

[180] *See* Mark Alan Thurmon, *The Rise and Fall of Trademark Law's Functionality Doctrine*, 56 FLA. L. REV. 243, 253-319 (2004).

[181] 514 U.S. 159 (1995).

[182] *Id.* at 165 (quoting Inwood Labs., Inc. v. Ives Labs. Inc. 456 U.S. 844, 850 n.10 (1982)).

enjoining the defendant's concurrent use of the feature.[183] Note also that, unlike misuse, fair use, and merger, functionality is not a defense. Instead, *nonfunctionality* is an element that the trade dress owner must prove.[184]

Commentators have identified as many as four distinct rationales that the functionality doctrine could, in theory, serve. The first, and broadest, would interpret the functionality doctrine to deny trademark protection to any product features that *could* be the subject of utility or design patent protection, on the theory that trademark protection for these features contravenes a federal "right to copy" that is implicit in cases such as *Bonito Boats, Inc. v. Thunder Craft Boats, Inc.*[185] A second rationale would interpret the functionality doctrine as avoiding a potential conflict between statutory regimes, if trademark law in effect would enable inventors to exercise exclusive rights in their inventions without requiring compliance with the more rigorous conditions imposed by patent law.[186] A third rationale posits that according exclusive rights in product designs that have been disclosed in utility patents postpones indefinitely the falling into the public domain of those features, thus violating the terms of the "patent bargain" that promises

---

[183] *See* Harold Weinberg, *Trademark Law, Functional Design Features, and the Trouble with* TrafFix, 9 J. INTELL. PROP. L. 1, 59-60 (2001).

[184] *See* 15 U.S.C. § 1125(a)(3) (2000); Tumblebus Inc. v. Cranmer, 399 F.3d 754, 768 (6th Cir.), *pet'n for cert. filed*, 73 U.S.L.W. 3734 (U.S. Jun. 13, 2005) (No. 04-1684). As for the other defenses, see Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 590 (1994) (fair use); A & M Records v. Napster, Inc., 239 F.3d 1004, 1026-27 (9th Cir. 2001) (copyright misuse); *Va. Panel*, 133 F.3d at 868 (patent misuse); *Kregos*, 937 F.2d at 705 (merger). *But see* Sony Corp. v. Universal City Studios, 464 U.S. 417, 451 (1984) (apparently shifting the burden of rebutting fair use to plaintiffs, upon proof that defendants' use was noncommercial).

[185] 489 U.S. 141 (1989). For articulations of this argument, see Margreth Barrett, *Consolidating the Diffuse Paths to Trade Dress Functionality: Encountering* TrafFix *on the Way to* Sears, 61 WASH. & LEE L. REV. 79, 96-104 (2004); Thurmon, *supra* note ___, at 343-44, 357.

[186] *See* Barrett, *supra* note ___, at 146-51.

exclusive rights for only a limited time.[187] A fourth rationale interprets functionality as precluding the assertion of trademark rights only when competitors *must* have access to a particular design feature in order to compete effectively.[188]

This last-mentioned "competitive-need" rationale differs from the three patent-related rationales insofar as it focuses solely upon whether, in the language of *Qualitex*, "exclusive use . . . would put competitors at a significant non-reputation-related disadvantage,"[189] and not at all upon perceived conflicts with patent law or policy. As a result, application of the competitive-need rationale alone can result in classifying some trade dress as "functional," even though the dress would appear nonfunctional from the standpoint of some or all of the patent-based policies; and in classifying other trade dress as "nonfunctional," even though it would be functional from the perspective of the patent-based policies. The first point is illustrated by the so-called aesthetic functionality doctrine, which applies to design features that bear no relation to the actual performance of a product, but which "confer[ ] a significant benefit that cannot practically be duplicated by the use of alternative designs."[190] For example, in in *Wallace International Silversmiths, Inc. v. Godinger Silver Art Co.*,[191] the Second Circuit affirmed a judgment that the plaintiff's baroque-patterned silverware design was aesthetically functional, on the ground that exclusive rights in the design would have enabled the plaintiff to "exclude competitors from using those baroque design elements necessary to compete in the

[187] *See id.* at 153-57; Thurmon, *supra* note ___, at 343, 370.

[188] *See* Barrett, *supra* note ___, at 151-53; Thurmon, *supra* note ___, at 282.

[189] *Qualitex*, 514 U.S. at 165 (citation omitted).

[190] RESTATEMENT (THIRD) OF UNFAIR COMPETITION § 17 cmt. c (1995).

[191] 916 F.2d 76 (2d Cir. 1990).

market for baroque silverware."[192] Because the design did not cause the silverware to perform any better than other designs, however, exclusive rights would not have created any conflict with the law of utility patents, and thus the design would not have been functional if only the utility patent-based rationales governed.[193]

To illustrate the second set of cases, in which competitive need would permit the exercise of trade dress rights but the patent-based rationales would not, consider a design feature X that is claimed in a utility patent as improving the quality of product Y when it is incorporated into that product. An advocate of the patent-based rationales would argue that, once the patent expires, product feature X must fall into the public domain and may not be the subject of trade dress protection--no matter how distinctive the feature is, and regardless of whether consumers rely upon feature X as a source-identifier. The competitive-need rationale, by contrast, would forbid the party marketing feature X from asserting trade dress rights *only* if the party's competitors must have access to feature X to compete effectively. Given that many patents do not confer market power, however, it is hardly a forgone conclusion that according exclusive rights to feature X *would* put competitors at a disadvantage. From an economic standpoint, whether or not there is a competitive need for feature X (or product Y) depends upon whether there are adequate substitutes, not upon the existence of a utility patent. Considerations of this nature led

[192] *Id.* at 81. Nevertheless, the court did refuse to allow competitors to market an *exact* duplicate of the plaintiff's design, in order to avoid confusion. *See id.* at 80.

[193] Some commentators nevertheless would condemn the exercise of exclusive trademark rights in purely aesthetic product features, on the ground that these rights do create a conflict with the law of *design* patents. *See* Barrett, *supra* note ___, at 151 n.305.

some courts and commentators to conclude that patent-based rationales express an empty formalism, and that the touchstone of functionality should be competitive need alone.[194]

The Supreme Court nevertheless disagreed with this view in its 2001 decision in *TrafFix Devices, Inc. v. Marketing Displays, Inc.*[195] In *TrafFix*, the Court held that product features that are essential to the use or purpose of an article, or that affect their cost or quality, are functional *regardless* of competitive need.[196] By way of justification, the Court noted that trade dress protection can be procompetitive, to the extent it dispels confusion, but that "protection must subsist with the recognition that in many instances there is no prohibition against copying goods and products," because such copying itself can promote competition and technological advance; and that courts therefore should strive to avoid "misuse or over-extension of trade dress."[197] Apropos of this latter point, the Court cited its decision from the preceding term in *Wal-Mart*,[198] in which the Court had cautioned against the abuse of trade dress litigation as a means of stifling legitimate competition.[199] As for that part of the *Qualitex* definition that refers to "significant non-

[194] *See* Thurmon, *supra* note ___, at 282-302 (discussing the prevalence of the competitive need rationale in the case law up to 2001, and arguing in favor of that standard); Thomas F. Cotter, *Is This Conflict Really Necessary? Resolving an Ostensible Conflict Between Patent Law and Federal Trademark Law*, 3 MARQUETTE INTELL. PROP. L. REV. 25, 39, 62-63 (1999).

[195] 532 U.S. 23 (2001).

[196] *See id.* at 32-33.

[197] *See id.* at 29.

[198] 529 U.S. 205 (2000).

[199] *See id.* at 213.

reputation-related disadvantage," the *TrafFix* Court asserted that it remains proper to consider competitive necessity in cases involving aesthetic functionality.[200]

Like the cases interpreting the merger doctrine, the cases addressing functionality provide little guidance with respect to the definition of key terms. As Robert Bone has noted, in determining whether a given design is "essential to the use or purpose of an article," judges are largely free to define the "article" however broadly or narrowly they want.[201] If the article is defined broadly enough—e.g., as "chairs" or "silverware"—there may be relatively few features that literally are essential (and a correspondingly large number of alternatives available to competitors, to the extent that competitive need remains a valid concern at least in aesthetic functionality cases). A narrow definition, limited to the specific product the plaintiff markets (e.g., Eames chairs, or baroque-patterned silverware) will often lead to the opposite result.[202] As Bone notes, however, "courts seldom bother to define the product before making the functionality determination. Instead, they seem to rely on intuition and make rough judgments, mostly implicit, about the appropriate level of generality at which to define the product."[203] To be sure, courts sometimes rely upon evidence, often in the form of expert testimony,

---

[200] *See TrafFix*, 532 U.S. at 32-33. Non-aesthetic functionality--the type at issue when a product feature is essential to the use or purpose of an article, or affects its cost or quality—is sometimes referred to as "utilitarian functionality." *See* RESTATEMENT (THIRD) OF UNFAIR COMPETITION § 17 cmt. c (1995).

[201] *See* Bone, *supra* note __, at 2174-75.

[202] *See id.*

[203] *See id.* at 2176; *see also* Kingsbury, *supra* note __, at 69-70. There appear to be no cases in which courts have applied antitrust standards to define the market affected by the exercise of trade dress rights, or to predict the impact upon competition within that market if the trade dress owner were accorded exclusive rights; though occasionally courts consider market shares. *See, e.g.*, Disc Golf Ass'n v. Champion Discs, Inc., 158 F.3d 1002, 1009 (9th Cir. 1998) (holding that design was functional, in light of evidence that 77% of disc golf courses used that design); *Mark Bric Display Corp. v. Joseph Struhl Co.*, No. C.A. 98-532ML, 2003 WL 21696318, at *6 (D.R.I. July 9, 2003).

concerning the number of comparable alternative designs open to competitors, as recommended by the *Restatement (Third) of Unfair Competition.*[204] When only a "limited number" of alternatives are available, the design is likely to be functional.[205] As should be familiar by now, however, deciding what constitutes a "limited number," or what makes one design sufficiently distinct from another, or whether access to a particular design is necessary to attain effective competition, rarely if ever involves detailed analysis of the economic consequences within any well-defined market.

As for the cases involving utilitarian functionality in particular, the rule that product features are functional if they are essential to the use or purpose of an article, or affect its cost or quality, without regard to competitive need is consistent with the patent-based rationales catalogued above. In conformity with the terminology I have used throughout this Article, one might say that to permit trade dress owners to assert exclusive rights in patentable product features would enable them to expand the scope of their grants, insofar as trademark law would enable the exercise of exclusive rights indefinitely. As I have suggested throughout, however, beyond-the-scope rationales such as this beg the question of what the appropriate scope *is*, and why. In cases such as *TrafFix* and *Wal-Mart*, the Supreme Court seems to be suggesting that the appropriate scope of trademark protection is narrow because otherwise the social costs, including the

[204] *See* RESTATEMENT (THIRD) OF UNFAIR COMPETITION § 17 cmt. b (1995); *see also* Thurmon, *supra* note ___, at 282-302. Although consideration of alternatives is more relevant to aesthetic than to utilitarian functionality after *TrafFix*, the existence of alternatives might also shed some light on whether a particular feature is in fact "essential" to the use or purpose of the article. On this basis some courts continue to take alternatives into account in for purposes of utilitarian functionality as well. *See* 1 MCCARTHY, *supra* note __, § 7.75 (arguing in favor of this approach); *cf.* Barrett, *supra* note ___, at 129-32 (disagreeing).

[205] *See* RESTATEMENT (THIRD) OF UNFAIR COMPETITION § 17 cmt. b (1995) (stating that "a design maybe functional if it is one of a limited number of superior designs"); *see also Disc Golf*, 158 F.3d at 1009; Chrysler Corp. v. Vanzant, 44 F. Supp. 2d 1062, 1072 (C.D. Cal. 1999) (holding that design was functional, in light of only three design alternatives); In re Lincoln Diagnostics, 30 U.S.P.Q.2d (BNA) 1817, 1824 (T.T.A.B.) (holding that seven alternatives were insufficient), *aff'd mem.*, 41 F.3d 1519 (Fed. Cir. 1994).

potential for trademark rights to interfere with competition and technological progress, outweigh the social benefits.[206] I will argue below that this assumption may be correct, but that it nevertheless is useful to develop a framework for comparing the relative costs and benefits. Moreover, if the Supreme Court is right, it also may be useful to consider whether product design trade dress protection, for either utilitarian or aesthetic features, is *ever* worth even a small risk of anticompetitive harm.

IV. Analyzing the Costs and Benefits of Alleged Anticompetitive Conduct

As the case law presented above shows, on some occasions antitrust and IP law diverge in their approach to the possibility that anticompetitive harm may flow from certain practices relating to patents, copyrights, and trademarks. To some extent, these differences may be arbitrary, or the result of historical accident. To the extent there are *good* reasons for this divergence, however, those reasons may be related to the differing costs, including error costs, of ruling for plaintiffs and defendants in some antitrust and IP contexts. In this Part, I present a thought experiment based upon a stylized model of a tribunal's decisionmaking process, both to illustrate the preceding points and to suggest how the real-world decisionmaking process necessarily differs from the theoretical ideal. I then consider the various types of harms that would be relevant to this analysis, in both the antitrust and IP contexts.

A. A Thought Experiment

The model begins by positing that a tribunal is charged with judging either an antitrust claim or an IP claim implicating one of IP's procompetitive doctrines, in a manner that will maximize expected social welfare. To carry out this charge, the tribunal

[206] *See TrafFix*, 532 U.S. at 1259-60; *Wal-Mart*, 529 U.S. at 213-14.

must predict and compare two different harms. The first harm, H_1, is the potential social loss that may result from a ruling *in favor of the antitrust defendant or the IP plaintiff*. The second harm, H_2, is the social loss that may result from a ruling *in favor of the antitrust plaintiff or the IP defendant*. The presumption is that the tribunal will rule for the antitrust defendant or IP plaintiff if the expected value of H_2 ($E(H_2)$) exceeds the expected value of H_1 ($E(H_1)$); it will rule in favor of the antitrust plaintiff or IP defendant if $E(H_1) > E(H_2)$. To maximize expected social welfare, however, the tribunal also must consider the cost of adjudication, the extent to which it can be confident that its estimates of H_1 and H_2 are accurate, and the relation between cost and greater accuracy.

The first step is to delineate the various components of H_1 and H_2. A major component of H_1 is the deadweight loss from monopoly, although I will argue below that in the IP context there are other components to consider as well. As for H_2, in the antitrust context this harm is composed of whatever cognizable efficiency losses (e.g., reduction of transaction costs) may result from condemning the conduct at issue.[207] In the context of the IP doctrines, on the other hand, an additional component of H_2 is the harm to the incentive structure of the IP system. A ruling in favor of the IP defendant, for example, on a fair use or misuse defense reduces the expected payoff from producing creative subject matter, and therefore may reduce the incentive of others to create and publish. This harm is certainly difficult to quantify—and for all we know, it may be relatively small across a wide range of endeavors—but the IP system is based in large part upon the assumption that it exists, and that IP rights are a corrective to the problem.

[207] Other components of H_2 in the antitrust context may include the private costs of structuring transactions so as to avoid potential liability, and the social costs of encouraging alternative transactions which themselves may give rise to the possibility of anticompetitive behavior, such as some forms of vertical integration. *See* Lopatka & Page, *supra* note __, at 387.

The tribunal therefore cannot ignore this harm, although as we shall see it may be able to discern some specific instances in which there is a high likelihood that the harm is small.

Second, the model assumes that the tribunal can either make a preliminary determination, based upon an initial submission of evidence, to rule in favor of the plaintiff or defendant; or it can defer its decision pending further evidence. Because the future is uncertain, however, the best the tribunal can do at either stage is to estimate probable future harm. In theory, one way to make this estimate would be for the tribunal to take a random sample of estimates of future harms that may flow from ruling one way or the other. For example, the tribunal could survey a panel of independent experts to obtain different assessments of the values of H_1 and H_2. Those assessments could be pooled to produce individual estimates of the means of H_1 and H_2. Alternatively, the tribunal could assign different weights to its own individual or collective estimates of these possible future consequences.[208] In theory, the tribunal could then use statistical techniques to determine if the sample data show conclusively that one mean is greater than another.

To illustrate, suppose first that the tribunal obtains five independent estimates of H_1, specifically (7, 4, 4, 10, and 15), denominated in some relevant monetary unit; and that it obtains five observations of H_2, specifically (12, 20, 17, 30, and 21). On the

[208] This is not necessarily as strange as it may sound. Real-world tribunals may be able to draw upon different the testimony of multiple expert witnesses (most of whom, however, will be retained by the parties and therefore may be biased). In addition, one can imagine a judge thinking, for example, that one way the future may play out is that, say, the user interface at issue will become an industry standard. If she rules against the IP defendant on its fair use, misuse, or merger defense, there will be a substantial deadweight loss. Alternatively, there is a possibility that the future may play out differently, with the interface quickly being overtaken by a competing standard, in which case the deadweight loss will be low. The judge could assign different probabilities to these alternative states of the world. Similarly, if the judge rules for the defendant, one possibility is a substantial weakening of the copyright incentive. Alternatively, potential creators and publishers might not be affected very much, due to the stringent conditions under which the IP doctrine at issue applies and in light of other means for reducing the free rider problem.

assumption that H_1 and H_2 are normally distributed, a t-test reveals an observed t-statistic of t = 3.33, leading to the statistically significant conclusion that $E(H_2) > E(H_1)$.[209] Abstracting from the cost of making the initial determination, the result is clear: the tribunal should rule in favor of the antitrust defendant or IP plaintiff. On the other hand, if one is not willing to assume that H_1 and H_2 are normally distributed, a non-parametric Mann-Whitney test could be used instead; it too would reveal a statistically significant conclusion that $E(H_2) > E(H_1)$.[210]

Alternatively, suppose that the estimates of H_1 and H_2 are (10, 1, 12, 0, 17) and (30, 4, 36, 19, 11), respectively. Repeating the t-test, the tribunal would find that t = 1.78 with a p-value of 0.11.[211] A similar result is found from a Mann-Whitney test, where the z-value is 1.57 and the p-value is 0.12.[212] Using either test, the difference between the H_1 and H_2 data is statistically insignificant. Notice that the means of the estimates, 8 and 20 respectively, are the same as in the first example; the greater variances in the second example preclude the inference of a distinction between the two harms. In a case like the latter one, then, the tribunal must choose from among three imperfect options. The first, and theoretically the best, would be to continue sampling until it obtained a statistically significant result. Continued sampling, however, most likely comes at some cost, and

---

[209] *See* MICHAEL O. FINKELSTEIN & BRUCE LEVIN, STATISTICS FOR LAWYERS 224 (2d ed. 2001). Using the hypothetical observations from the above text, the t-score comes to -3.33, which corresponds to a p-value of 0.01. *See id.* at 571 tbl. E. In many scientific studies, a p-value of .05 or less is considered to be statistically significant.

[210] To employ the Mann-Whitney test, one would rank the ten observations (five for H_1 and five for H_2) as follows: 4, 4, 7, 10, 12, 15, 17, 20, 21, 30. The sum of the ranks (S) of the five H_1 observations (1 + 2 + 3 + 4 + 6 =16) is statistically significant. *See id.* at 343-44 & 580 tbl. H2. The related U-statistic is 24, *see id.*, which corresponds to a z-value of 2.40 and a p of 0.02 (though it would be preferable to compute the z-value on the basis of a larger sample size). *See id.* at 557 tbl. A2; SHARON WEINBERG & KENNETH GOLDBERG, STATISTICS FOR BEHAVIORAL SCIENCE 495-96 (1990).

[211] *See* FINKELSTEIN & LEVIN, *supra* note ___, at 571 tbl. E.

[212] *See id.* at 557 tbl. A2; WEINBERG & GOLDBERG, *supra* note __, at 495-96.

thus the tribunal would have to make some rough prediction whether the expected cost of additional sampling would exceed the expected value of the sampling; this in turn will depend upon the stakes involved (that is, the magnitude of H_1 and H_2), how much more sampling the tribunal expects would be necessary to obtain a statistically valid result (and how would it know that?),[213] and the expected cost of such additional sampling. A second option would be to rule on the basis of whether $E(H_1) > E(H_2)$ or $E(H_1) < E(H_2)$, even though the difference between these means is statistically insignificant, on the ground that this is simply the best that can be done under the circumstances. A third option, again assuming that further investigation is not feasible, would be to rule in favor of the antitrust defendant whenever the result of the initial determination is statistically insignificant. In doctrinal terms, the tribunal would conclude in such a case that the antitrust plaintiff has not met its burden of proof.[214] In the IP context, on the other hand,

---

[213] *Cf.* JOHN KAY, CULTURE AND PROSPERITY: WHY SOME NATIONS ARE RICH BUT MOST REMAIN POOR 219-20 (2004) ("How could the economist know when to stop calculating when he cannot know the benefits of further calculation?"); Beckner & Salop, *supra* note __, at 45-47.

[214] Note, however, that the relationship between statistical significance and the legal burden of proof is not a simple one. In litigation in which scientific evidence is proffered, courts are reluctant to admit evidence that fails to meet some definition of statistical significance. For example, in a civil action in which the plaintiff alleges that exposure to the defendant's product caused her to incur cancer, the court is likely to admit, as evidence relevant to causation, only studies indicating at the 90% or 95% significance level that exposure to the product increases the risk of cancer. In general, an epidemiologist would not rely upon studies with higher false positive rates as evidence that the substance causes cancer, and a court is similarly unlikely to view the study as sufficiently reliable to assist the trier of fact in determining the cause of the plaintiff's cancer. *See, e.g.*, Merrell Dow Pharmaceuticals, Inc. v. Havner, 953 S.W.2d 706, 724 (Tex. 1997); Susan R. Poulter, *Science and Toxic Torts: Is There a Rational Solution to the Problem of Causation?*, 7 HIGH TECH. L.J. 189, 261-62 (1992); *cf.* Michael D. Green, *Expert Witnesses and Sufficiency of Evidence in Toxic Substances Litigation: The Legacy of* Agent Orange *and the Bendectin Litigation*, 86 NW. U. L. REV. 643, 685-91 (1992) (arguing that "rejecting all studies that are not statistically significant would be a cursory and foolish judgment, particularly if there are multiple studies tending to show a consistent effect," and that the better practice would be to admit some such studies where the magnitude of the association, and the risk of false negatives, are sufficiently great). Applying a restrictive causation standard, however, does not mean that the court has increased the plaintiff's burden of proving causation above a "mere preponderance" (50%). *See* Michael D. Green, *Science Is to Law as the Burden of Proof Is to Significance Testing*, 37 JURIMETRICS J. 205, 221-22 (1997) (reviewing CARL F. CRANOR, REGULATING TOXIC SUBSTANCES: A PHILOSOPHY OF SCIENCE AND THE LAW (1993)); David H. Kaye, *Apples and Oranges: Confidence Coefficients and the Burden of Persuasion*, 73 CORNELL L. REV. 54, 65-66 (1987).

the default rule in an ambiguous case may depend upon whether the doctrine at issue is classified as an affirmative defense, with respect to which the defendant bears the burden of proof, or an element of the plaintiff's case-in-chief. As noted above,[215] under current law, misuse is an affirmative defense, as is fair use and merger. *Non*functionality, on the other hand, is in general an element of the trademark plaintiff's case, which suggests that in an ambiguous case the court should rule for the trademark defendant.

For obvious reasons, a model like the preceding one differs from real-world practice in many particulars. Nevertheless, in its broad contours it captures some important real-world elements. The first is that, although a court is unlikely to engage in multiple random sampling of future events as such, it can and should engage in some analysis of how the various possible ways the future may play out if it rules for one side or the other. The analysis may be rough but is clearly consistent with antitrust jurisprudence, which counsels courts to consider the pro- and anticompetitive effects of contested transactions and practices, and with those IP doctrines that express concern for competition. Second, it is unlikely that real-life tribunals can make anything more than an educated guess about the actual magnitudes of H_1 and H_2, whether at the preliminary or a later state of the proceedings, or about the costs and benefits of further investigation concerning those harms. What may be possible, however, is to engage in considered analysis of what those harms are likely to consist of and which harm is likely to dominate in a given case. For example, there may be cases in which the court can conclude that the probability of anticompetitive harm is relatively low but the *expected* harm, in the sense

[215] *See supra* notes __ and accompanying text.

of probability multiplied by magnitude if the event occurs, is high; even this approximate comparison may be useful, as I discuss below.

Finally, note that one can relate the preceding analysis to the harms associated with false positives and false negatives. To see why, consider the consequences of a court ruling incorrectly, that is, of ruling for the antitrust defendant/IP plaintiff in a case in which H_1 winds up exceeding H_2, or in favor of the antitrust plaintiff/IP defendant in a case in which H_2 winds up exceeding H_1. The excess of harm flowing from an incorrect ruling in favor of the antitrust defendant/IP plaintiff can be thought of as the harm attributable to a false negative, whereas the excess of harm flowing from an incorrect ruling in favor of the antitrust plaintiff/IP defendant can be thought of as the harm attributable to a false positive. I will suggest below that there may be some instances in which the expected cost of false positives is predictably lower in the IP than in the antitrust context, and the expected costs from false negatives predictably higher. If so, this provides a rationale, in an appropriate case, for excusing IP defendants from liability on the basis of harms that would be unduly speculative or non-cognizable in antitrust.

B. H_1 and H_2 in Antitrust Law

In assessing whether $E(H_1) > E(H_2)$ in the antitrust context, a court must engage in careful factual analysis. First, it must be reasonably sure that H_1 is real, not merely some phantom harm. To be sure, in some cases involving per se liability, such as naked price fixing, the court may dispense with the need to define the market and prove market power resulting in deadweight loss. This strategy makes sense as a means for reducing adjudication costs, because in the absence of power over a definable market antitrust defendants normally would have no reason to enter into such an agreement in the first

place. In most other cases, however, market definition and proof of market power are required to ensure that the potential H_1 harm is real. As a result, one of the more contentious issues in antitrust today involves the concept of harm to so-called "innovation markets," in accordance with which a transaction or practice is subject to challenge on the ground that it may inhibit future innovation.[216] The government's case against Microsoft, for example, involved allegations that Microsoft's business practices threatened to inhibit innovation in the software and middleware markets;[217] and on occasion the FTC has considered the effect on innovation markets in deciding whether to challenge horizontal mergers.[218] Concern over innovation markets in the antitrust context

[216] According to the IP Guidelines:

> A licensing arrangement may have competitive effects on innovation that cannot be adequately addressed through the analysis of goods or technology markets. For example, the arrangement may affect the development of goods that do not yet exist. Alternatively, the arrangement may affect the development of new or improved goods or processes in geographic markets where there is no actual or likely potential competition in the relevant goods.

IP GUIDELINES § 3.2.2, at 10-11. The DOJ and the FTC have asserted potential harm to innovation markets as a basis for proceeding with actions against both Microsoft and Intel, among others, and for imposing conditions for the approval of some mergers. *See* Davis, *supra* note __, at 687-94; Richard J. Gilbert & Willard K. Tom, *Is Innovation King at the Antitrust Agencies? The Intellectual Property Guidelines Five Years Later*, 69 ANTITRUST L.J. 43, 47-82 (2001). Defenders of the innovation markets concept argue that the potential harm at issue is substantial, and that traditional antitrust doctrines are not sufficient to protect against it. *See* Richard J. Gilbert & Steven C. Sunshine, *Incorporating Dynamic Efficiency Concerns in Merger Analysis: The Use of Innovation Markets*, 63 ANTITRUST L.J. 569, 587 (1995); Tom & Newberg, *supra* note __, at 222-28. Critics contend that the harm is unduly speculative, given how little we know about the optimal conditions for inducing innovation, and that conventional antitrust tools can deal with provable harm. *See* Lopatka & Page, *supra* note __, at 370-72; Richard T. Rapp, *The Misapplication of the Innovation Market Approach to Merger Analysis*, 64 ANTITRUST L.J. 19, 37-46 (1995).

[217] *See* Brief for Appellees United States and State Plaintiffs at 16-20, United States v. Microsoft Corp., 253 F.3d 34 (D.C. Cir. 2001) (Nos. 00-5212, 00-5213); Brief for United States at 31, United States v. Microsoft Corp., 373 F.3d 1199 (D.C. Cir. 2003) (No. 03-5030).

[218] *See* Davis, *supra* note __, at 687-94.

is nevertheless controversial precisely because H_1 is relatively speculative, in comparison with the harm at issue in a typical antitrust case.[219]

At the same time, antitrust courts must pay close attention to the potential harm from false positives. In evaluating conduct that has both pro- and anticompetitive potential, courts must consider both the frequency of potential errors and the magnitude of those errors. The latter can be substantial, not only because the harm resulting from a decrease in allocative efficiency is serious in and of itself, but also because various aspects of antitrust law tend to magnify the impact of that harm.[220] For example, courts automatically treble antitrust damages and award attorneys' fees to the successful plaintiff.[221] Treble damages awards in particular raise a risk of overdeterring firms from engaging in conduct that promises some efficiency gains but that may be difficult for a court to distinguish from anticompetitive conduct.[222] Other potential sources of overdeterrence include the potential for criminal penalties,[223] though typically these are encountered only in connection with offenses such as price fixing; additional civil sanctions, such as the risk of indirect purchaser suits under state law;[224] and other costs

---

[219] *See supra* note ___.

[220] *See* Lopatka & Page, *supra* note __, at 387.

[221] *See* 15 U.S.C. § 15.

[222] *See* Lopatka & Page, *supra* note __, at 387; *cf.* Thomas F. Cotter, *An Economic Analysis of Enhanced Damages and Attorneys' Fees for Willful Patent Infringement*, 14 FED. B.J. 291, 318-26 (2004) (discussing how damages multipliers may have a similar overdeterrent effect in patent litigation).

[223] *See* 15 U.S.C. §§ 1, 2.

[224] In a case in which the defendant is found to have engaged in anticompetitive conduct, such that the measure of damages before trebling is the amount of the monopoly overcharge, a party that purchases directly from the defendant may be able to pass some or all of that overcharge on to its customers, and so on down the chain of distribution. Federal antitrust law nevertheless forbids courts from awarding antitrust damages to indirect purchasers, out of concern for, among other things, the potential for overlapping or inconsistent damages awards. *See* Ill. Brick Co. v. Illinois, 431 U.S. 720, 730-31 (1977). State antitrust laws, however, often permit indirect purchasers to pursue their own damages awards, thus raising a

including the impact upon management time and the price of company shares.[225] Yet another concern, especially in cases involving unilateral conduct, is the social cost of having courts or agencies monitor compliance with court orders.[226] Finally, in cases involving the alleged anticompetitive use of IP rights, there is an additional risk that an aggressive antitrust policy might have a negative impact on innovation, by decreasing the expected returns from the exercise of IP rights. Thus, even if $E(H_1)$ appears to exceed $E(H_2)$, an antitrust court arguably should be reluctant to intervene unless it is confident that the risk of a false positive is low.

C. H_1 and H_2 in IP Law

In considering whether $E(H_1) > E(H_2)$ in the IP context, a court similarly must engage in careful analysis of the relevant harms. Consider first the potential H_2 harms. As in the antitrust context, H_2 may include static efficiency losses. For example, in cases involving allegations of patent or copyright misuse, a ruling for the IP defendant may deprive the parties of efficient means of reducing transaction costs (such as pools or package licensing) or of financing (such as by extending the royalty term past the term of protection). But not all IP cases involve potential static efficiency losses. It is difficult to perceive, for example, what short-term efficiency losses would result from a ruling for the IP defendant in a case such as *Sega* or *Sony*, in which the IP plaintiff asserts the

---

potential overdeterrence problem. *See* William H. Page, *Class Certification in the* Microsoft *Indirect Purchaser Litigation*, at 4, 47, *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=671048.

[225] *See* Lopatka & Page, *supra* note __, at 387; *see also* Meurer, *supra* note __, at 1875 (arguing that antitrust presents greater incentives than does IP law for "opportunistic or anticompetitive litigation").

[226] Even outside the IP context, courts are often unsympathetic to claims that a monopolist's unilateral refusal to deal violates Sherman Act § 2, partially out of concern that a ruling for the antitrust plaintiff may require the imposition and subsequent supervision of the terms of an ongoing business relationship. *See* POSNER, *supra* note __, at 242.

unilateral right to forbid others from copying its software for the purpose of reverse engineering. A ruling for the defendant should have, if anything, a net *positive* effect on static efficiency, by eliminating any deadweight loss (H_1) attributable to the IP; and, because a judgment excusing the IP defendant from liability effectively awards that defendant a royalty-free license, the cost of monitoring compliance typically should be lower than in many antitrust settings.

There are, however, other components of H_2 that must be considered in the IP context; most prominent is the potential harm to dynamic efficiency. Weakening the IP right in one case reduces others' ex ante incentives to create, publish, or invest in goodwill, to some degree, and this reduction in incentives may mean that future consumers will have fewer choices than they otherwise would have.[227] This potential harm is at least theoretically present in every IP case; and while it may be unquantifiable, the IP system is largely premised upon the assumption (right or wrong) that this harm is, in general, substantial. Otherwise there would be no need for IP rights in the first place, at least not from an instrumental perspective. Policymakers also should take into consideration other possible future harms--for example, that IP owners will try to circumvent application of a procompetitive doctrine (such as a ruling that reverse engineering constitutes fair use) by adopting practices that evade condemnation but which raise other costs (such as making their products more difficult to reverse

[227] Indeed, misuse doctrine may exhibit an alarming potential to undermine incentives, to the extent it renders a patent or copyright unenforceable in its entirety until the misuse is purged, and not merely unenforceable with respect to a specific use. On the other hand, a logical response to misuse caselaw may be for IP owners to structure their transactions differently in the future, rather than to stop creating. In such instances, the potential efficiency loss, whether viewed as short- or long-run, may consist largely of higher transactions costs; though to the extent a forbidden license term embodies an effective way of enforcing one's IP rights, misuse doctrine may reduce payoffs and hence incentives too.

engineer).[228] Ideally, the cost incurred to make the product more difficult to crack, as well as the cost incurred by reverse engineers in response, would be weighed in comparison with whatever cost reductions reverse engineering gives rise to.[229]

Nevertheless, the assumption that the dynamic-efficiency component of H_2 is *in general* substantial does not imply that this component is *always* substantial. If a tribunal can be confident that the expected value of H_2 (including both the static and dynamic components) in a specific case is low (or at least lower than the expected value of H_1), excusing the IP defendant from liability would appear to enhance social welfare. To be sure, the tribunal should be extremely careful in making such assessments so as to avoid costly false positives. It is often claimed, for example, that only a small percentage of the money invested in pharmaceutical research results in the development of a commercially successful drug (though how much actually is invested remains an open question).[230] In considering the value of the patent incentive, then, one needs to take into account the number of "dry holes" from which no patentable invention was forthcoming.[231] Another caveat is the possibility of a so-called "lottery effect," i,e., that some inventors and creators really are motivated by the small ex ante probability of earning vast rewards from their creations; cutting off those rewards because they appear excessive ex post may

[228] *See infra* notes __ and accompanying text.

[229] *See* Samuelson & Scotchmer, *supra* note __, at 1625. Alternatively, as explained below, reverse engineering may render IP owners' attempts to price-discriminate ineffective. *See infra* note __.

[230] *See, e.g.*, Dan L. Burk & Mark A. Lemley, *Policy Levers in Patent Law*, 89 VA. L. REV. 1575, 1616 (2003).

[231] *See* Katz, *supra* note __, at 20-21.

have the unintended consequence of deterring risky innovation.[232] Even so, scholars have analyzed a handful of settings in which the ex post return on investment may well be sufficient to induce comparable investment on the part of others, even if legal doctrine caps that return to some extent.[233] *If* courts can identify cases exhibiting these characteristics with reasonable confidence, both the frequency and magnitude of false positives associated with harm to dynamic efficiency should be low. They include the following.

Cases in which network effects are present. A network effect is said to exist whenever "the utility that a user derives from consumption of a good increases with the number of other agents consuming the good."[234] The most direct example is the telephone, which is useless if only one user is connected to the phone network,[235] but network effects also may be generated by other products or services that become industry standards. Many aspects of computer technology, for example, including operating systems and some applications programs and applications program interfaces, may

[232] *See* Scherer, *supra* note __, at 19-21. In many institutional settings, however, managers may be more risk-averse than the risk-loving individuals considered by Scherer.

[233] This assumes, however, that other potential creators and producers perceive the exception as being applied in such a way that they have little to fear if they are subjected to similar treatment in the future. To the extent an exception is perceived as having been wrongly or inconsistently applied, there is a greater risk of impact upon future incentives. See James A. D. White, *Misuse or Fair Use: That Is the Software Copyright Question*, 12 BERKELEY TECH. L.J. 251, 279 (1997) (arguing that, in light of the uncertainty that may arise from applying IP doctrines so as to adjust the "scope of the grant," courts should try to "balance the rights for a class of works rather than for each individual work"). In the text above, I attempt to isolate recurring situations in which both the actual and perceived frequency and magnitude of false positive errors should be minimal.

[234] Mark A. Lemley & David McGowan, *Legal Implications of Network Economic Effects*, 86 CAL. L. REV. 479, 483 (1998) (quoting Michael L. Katz & Carl Shapiro, *Network Externalities, Competition, and Compatibility*, 75 AM. ECON. REV. 424, 424 (1985)).

[235] *See id.* at 488-91.

exhibit network effects.[236] Of course, the presence of network effects with respect to a product that embodies IP rights is not a sufficient condition to ignore the IP rights. The potential for exploiting those rights may have been the necessary inducement for entering the risky standards competition that the plaintiff happened to win (and that other potential standard-bearers lost).[237] Because the rewards can be so disproportionately large, however—as can the potential for anticompetitive harm, which factors into the value of H_1 below[238]--the presence of network effects should be a relevant factor in computing the balance. All other things being equal, restricting the exercise of an IP right that, by the time the restriction is entered, already has earned the owner substantial returns, may not create a disincentive for others to invest in creating their own IP.[239]

Cases in which other incentives are present. In other instances, it may be reasonable to conclude that other incentives beside the IP rights at issue are sufficient to induce others to create, publish, or invest. For example, the amount invested in a specific

---

[236] *See, e.g.*, *id.* at 491-92; O'Rourke, *supra* note __, at 1179, 1212, 1217; Samuelson & Scotchmer, *supra* note __, at 1617; Weiser, *supra* note __, at 568-83.

[237] *See* Lemley & McGowan, *supra* note __, at 530; O'Rourke, *supra* note __, at 1217; Weiser, *supra* note __, at 586.

[238] *See infra* notes __ and accompanying text.

[239] *See* Lemley & McGowan, *supra* note __, at 530, 604; O'Rourke, supra note __, at 1217; Weiser, *supra* note __, at 591-92; White, *supra* note __, at 277-79. Depending upon the nature of the limitation, it might even increase others' incentives to create follow-up innovation. *See infra* notes __ and accompanying text. *Cf.* Rochelle Dreyfuss, *Unique Works/Unique Challenges at the Intellectual Property/Competition Law Interface*, *in* EUROPEAN COMPETITION LAW ANNUAL 2005: THE INTERSECTION BETWEEN COMPETITION LAW AND INTELLECTUAL PROPERTY LAW __, __ (Claus-Dieter Ehlermann & Isabela Atanasiu, eds., 2005) (arguing that competition law should take a more aggressive approach to ensuring access to "unique" works, especially where "the payoffs associated with the rights are extraordinarily high relative to the inventive contributions made" and "there is little danger that reducing the licensor's flexibility would strip it of the ability to earn an adequate return on investment"). Courts still need to be careful, however. Ensuring access to a technologically inferior standard may simply enable that standard to maintain its dominance; it may be better to promote competition among standards. *See* David McGowan, *Regulating Competition in the Information Age: Computer Software as an Essential Facility Under the Sherman Act*, 18 HASTINGS COMM. & ENT. L.J. 771, 848-49 (1996); Weiser, *supra* note __, at 590.

creative process may be so low that no significant free-rider problem exists at all, particularly if the creator has a first-mover advantage. Some minimally creative works, such as the sweepstakes instructions at issue in *Morrissey*,[240] might fall within this category.[241] In the patent arena, scholars have expressed doubt whether the patent incentive is necessary to induce the creation and disclosure of subject matter such as business methods.[242] Similarly, if other incentives such as direct or indirect government funding are present, the need for an additional boost from patent law may be small.[243] The presence of other IP protection also may be relevant. Reducing or even eliminating protection for trade dress, for example, may have little impact upon firms' incentives to invest in quality control, assuming that most products are sold in connection with other source identifiers such as word marks. Thus, even if others can copy my distinctive product design, I may have a sufficient incentive to produce goods of consistent quality, as long as I continue to market those goods under my brand name.[244]

Cases in which the term of the right has been extended retroactively. Yet another setting in which the dynamic component of H_2 might be small is when the IP plaintiff is the beneficiary of a windfall increase in the term of protection. As an example of this

---

[240] 379 F.2d 675 (1st Cir. 1967).

[241] Alternatively, authors and publishers of works they believe to be divinely inspired might have reason to disseminate those works, even in the absence of a copyright incentive. *See* Cotter, *supra* note __, at 363 (discussing disputes over copyright in purportedly scriptural works).

[242] *See, e.g.*, Rochelle Cooper Dreyfuss, *Are Business Method Patents Bad for Business?*, 16 COMP. & HIGH TECH. L.J. 263, 275 (2000).

[243] *See* Rebecca Eisenberg, *Public Research and Private Development: Patent and Technology Transfer in Government-Sponsored Research*, 82 VA. L. REV. 1663, 1668-70 (1996).

[244] To be sure, consumers may encounter the product without its brand name; this possibility is the rationale for protecting trade dress owners from the likelihood of post-sale confusion. *See* Nabisco, Inc. v. PF Brands, Inc., 191 F.3d 208, 218-19 (2d Cir. 1999); *see also infra* notes __ and accompanying text.

phenomenon, consider the Copyright Term Extension Act, which added twenty years to the copyright term both for future works and for existing works that were still under copyright protection as of 1998.[245] The likelihood that this extension serves a public purpose with respect to future works is slim, given the small present value of an additional twenty years' worth of protection in the distant future;[246] the likelihood that it serves a public purpose with respect to existing works is even slimmer, though theoretically possible.[247] Even with proper deference to Congress's contrary judgment that the extension serves a public purpose, however, a court rationally could conclude that the impact upon incentives of deciding a close fair use case, involving a work that is within the last twenty years of its term, in favor of the IP defendant is *de minimis*.[248] On the other hand, Congress generally authorizes patent term extensions only upon proof that the economic life of a patent has been shortened, due to regulatory or PTO delay.[249] For a court to apply the misuse doctrine more liberally during an extended patent term, on the

[245] Pub. L. No. 105-298, 112 Stat. 2827, 2827-28 (1998), *codified at* 17 U.S.C. §§ 301-04 (2000). *See also* William F. Patry & Richard A. Posner, *Fair Use and Statutory Reform in the Wake of* Eldred, 92 CAL. L. REV. 1639, 1646 (2004) (asserting that restricting windfalls impairs no incentive).

[246] *See* Eldred v. Ashcroft, 537 U.S. 186, 265-66 (2003) (Breyer, J., dissenting) (citing to Brief Amici Curiae George Akerlof et al.).

[247] In rejecting a constitutional challenge to the act, the Supreme Court accepted as possible justifications the possibilities that past authors may have been motivated to create in light of the possibility of future term extensions, and that term extensions may provide an incentive to republish or restore works which otherwise would fall into the public domain, forgotten and unrestored. *See Eldred*, 537 U.S. at 206-07.

[248] Some scholars have argued that the scope of fair use should expand as a work ages, citing among other reasons the attenuating effect on incentives. *See* Justin Hughes, *Fair Use Across Time*, 50 UCLA L. REV. 775, 800 (2003); Joseph P. Liu, *Copyright and Time: A Proposal*, 101 MICH. L. REV. 409, 433-46 (2002). Neither I nor these scholars are advocating that courts simply ignore Congress's judgment. But since fair use invites consideration of many factors in deciding whether to excuse a defendant from liability, it may not be unacceptable to consider among other things the waning impact on incentives over time.

[249] *See* 35 U.S.C. §§ 155-56 (2000).

theory that applying the doctrine in such a way will have no effect upon the relevant patent incentive, might be a more substantial intrusion upon congressional judgment.

Whatever the value of H_2 might be, the other side of the coin is H_1, and here too some IP cases may depart in significant ways from the typical antitrust case. In the IP context, H_1 would include deadweight loss whenever the IP at issue confers a measure of market power (which may, however, constitute only a subset of all IP cases). Even so, it is worth noting that in some cases involving IP rights the magnitude of H_1 is likely to be *quite* large. In particular, when network effects are present, the deadweight loss resulting from the exercise of the right may be substantial.[250] Moreover, network effects may facilitate the exercise of tying and other anticompetitive strategies which, in more run-of-the-mill cases, are unlikely to cause anticompetitive harm.[251] In addition, there may be reasons to include within H_1 a variety of harms that would not be cognizable, often or at all, in the antitrust context.[252] When the following harms are present, the mean value of H_1 may be greater than one might at first blush expect.

---

[250] *See supra* text accompanying notes ___. Even in cases not involving network effects, deadweight loss may be high, if the IP at issue occupies a distinct market. For arguments that IP rights often result in larger deadweight losses than contemporary scholars and regulators assume, see Katz, *supra* note __.

[251] In recent years, economists have identified situations in which it *may* be possible to leverage a monopoly in one product market into a monopoly in another market. *See, e.g.*, Dennis W. Carlton & Michael Waldman, *The Strategic Use of Tying to Preserve and Create Market Power in Evolving Industries*, 33 RAND J. ECON. 194 (2002) (presenting models in which a firm with monopoly power in the market for a primary good uses tying or bundling to exploit network externalities in the market for a complementary good, so as to preserve market power in the market for the primary good or in an emerging market for a new good); Jay Pil Choi, *Antitrust Analysis of Tying Arrangements*, CESIFO WORKING PAPER NO. 1336 (Nov. 2004), *available at* http://ssrn.com/abstract=629001 (reviewing literature suggesting, *inter alia*, that network effects can assist firms in preserving market power or extending into related markets). Absent the conditions identified in works such as these, however, the common view among economists today is that leveraging is *not* a likely result of tying. *See* KEITH N. HYLTON, ANTITRUST LAW: ECONOMIC THEORY & COMMON LAW EVOLUTION 281-83 (2003).

[252] *Cf.* Robin C. Feldman, *The Insufficiency of Antitrust Analysis for Patent Misuse*, 55 HASTINGS L.J. 399, 400 (2003) (arguing that patent misuse doctrine can take into account harms that antitrust does not recognize, including the social "waste that can occur with defensive research or inventing around a patent," "the burden on innovation that can result from an overproliferation of patent rights," and "the disincentives

Small, localized, or transitory competitive harms. As we have seen, antitrust law generally does not concern itself with competitive harms that are believed to fall below some threshold. For example, for purposes of defining a market, the courts and agencies typically focus upon whether a 5% price increase could be sustained for the foreseeable future.[253] As a result, contemporary antitrust law for the most part rejects the notion that courts should concern themselves with so-called "submarkets."[254] To be sure, the harm accruing from a 1% price increase for a short period of time is no less real than the harm accruing from a 5% increase; it's just that the harm is not deemed substantial enough to merit antitrust scrutiny, with its attendant adjudication costs and the ever-present risk of false positives. Even so, it does not necessarily follow that IP law should never be deployed as a counter to such relatively minor harm. In theory, *if* $E(H_2)$ is low enough, $E(H_1)$ may exceed $E(H_2)$ even though the value of $E(H_1)$ would not be "substantial" for antitrust purposes.[255] On the other hand, even if $E(H_1)$ appears to exceed $E(H_2)$, the

---

to innovation that can result from allocating reward to early-stage inventors over lat-stage inventors"). Although I agree with Professor Feldman that misuse doctrine sometimes should be more sensitive than antitrust to harm to innovation, I am less sanguine than she about an *expansive* role for misuse in combating the harms she lists. *See infra* Part V.A. More generally, though, Feldman makes the valid point that IP rights give rise to a variety of social costs, including but not limited to monopoly costs; and that whereas antitrust is directed only against the latter type of cost, IP doctrine rightly tries to minimize a wider class of costs. *See id.* at 438.

[253] *See* HORIZONTAL MERGER GUIDELINES, *supra* note __, § 1.11, at 7.

[254] Older antitrust cases sometimes refer to the potential for anticompetitive harm within submarkets, *see, e.g.*, Brown Shoe Co. v. United States, 370 U.S. 294, 325 (1962), and occasionally such language creeps into more recent case law as well. Most contemporary courts and scholars, however, take the view that "there are no submarkets, only markets." Jonathan B. Baker, *Stepping Out in an Old* Brown Shoe*: In Qualified Praise of Submarkets*, 68 ANTITRUST L.J. 203, 206 (2000); *see also* Geneva Pharmaceuticals Tech. Corp. v. Barr Labs. Inc., 386 F.3d 485, 496 (2d Cir. 2004). Even Baker, in his "qualified praise," actually defends only the uncontroversial proposition that well-defined markets may exist within larger, well-defined markets. *See* Baker, *supra*, at 207.

[255] *Cf.* Robert P. Merges, *Reflections on Current Legislation Affecting Patent Misuse*, 70 J. PAT. & TRADEMARK OFF. SOC'Y 793, 800 (1988) (stating, in support of the continued viability of the patent misuse doctrine, that "the often very limited (or 'thin') markets for patented technology make it difficult to apply antitrust law's consumer-demand definition of the relevant market"). Of course, one could ask why, in

decisionmaker may not sufficiently confident of this result; and the expected cost of further investigation may counsel against efforts to fine-tune the analysis. All other things being equal, in such a case the burden of proof may determine which side prevails.

Aggregate Competitive Harm. Proceeding from the above example, the case for exempting the IP defendant from liability may be stronger if $E(H_1)$, though small in the case at bar, is potentially large in the aggregate. To illustrate, suppose that there are five possible designs for a product that inhabits a discrete product market; or five equally marketable ways of expressing a given idea. According exclusive trademark or copyright rights to any one of the five alternatives might not cause much competitive harm, in and of itself, as long as the other four remain available to competitors. $E(H_1)$, viewed atomistically, is still small. Indeed, even if all five alternatives wind up in private hands, there may not be a substantial competitive effect; five independent firms may define a reasonably competitive market.[256] The risk of anticompetitive harm increases, however, the fewer the firms there are; and if access to at least one of the five alternatives *is* indispensable to competition within the relevant product market, the five firms effectively have created a barrier to new entry. Conceivably, even a relatively small probability of substantial aggregate H_1 harm might be sufficient to outweigh a low risk of H_2 harm. By contrast, unilateral conduct that does not threaten substantial competitive harm would rarely if ever be an appropriate subject of antitrust scrutiny, even if similar conduct in the

---

theory, antitrust too should not condemn conduct that threatens a small amount of anticompetitive harm, in a case in which the expected costs of a false positive is zero. One answer is that antitrust already may condemn such minimal harm in the subset of cases in which proof of market harm is not required, e.g., naked price fixing. Another is that antitrust enforcement carries with it numerous collateral social costs, as described above. In the IP context, by contrast, by hypothesis the plaintiff has already commenced litigation, and the question at issue is whether to enforce her rights.

[256] *Cf.* IP GUIDELINES, *supra* note __, § 4.3, at 23 (stating that the agencies normally will not challenge a restraint involving an innovation market if, *inter alia*, there are five independent firms within that market).

aggregate would result in social losses, as long as the firms engaging in the conduct are not expressly or implicitly colluding with one another.[257]

Harm to Innovation Markets. As noted above, an innovation market is a market for a product that does not yet exist but which may exist in the future, if innovation proceeds in some expected fashion.[258] Although the concept has been deployed in a smattering of antitrust cases, even the strongest defenders of the innovation market concept are chary about extending its reach too far. The potential for false positives and the magnitude of adjudication costs are simply too great to expect antitrust enforcers to make use of the innovation market concept under many circumstances.[259] A similarly cautious approach, however, may not be appropriate in some IP cases. For example, suppose that one could say with confidence that there is a 30% probability of incurring an H_1 harm of $10,000,000. $E(H_1)$ would be $3,000,000. If one also could be confident that $E(H_2)$ is less than $3,000,000, then the appropriate response would be to excuse the IP defendant from liability, despite the fact that the probability of incurring the H_1 harm is less than 50%. Particularly when network effects are present, this may well be a plausible outcome. In such a case, $E(H_1)$ is likely to be large in an absolute sense. In addition, the presence of network effects may make it plausible for the IP owner to successfully engage in conduct that otherwise would be unlikely to succeed, such as the

[257] *Cf.* Cargill, Inc. v. Monfort of Colorado, Inc., 479 U.S. 104, 125 (1986) (Stevens, J., dissenting) (arguing, unsuccessfully, that in evaluating horizontal mergers under § 7 of the Clayton Act, courts should consider "'the cumulative centripetal effect of acquisitions by large corporations, none of which by itself might be sufficient to constitute a violation of the Sherman Act'") (quoting *Brown Shoe*, 370 U.S. at 317).

[258] *See supra* notes __ and accompanying text.

[259] *See* 1 FEDERAL TRADE COMM'N, ANTICIPATING THE 21ST CENTURY: COMPETITION POLICY IN THE NEW HIGH-TECH, GLOBAL MARKETPLACE 5 (1996) (cautioning that "innovation market analysis is not always appropriate," and that "innovation market analysis should be applied only where the innovation is directed toward a particular good and where the innovation can be associated with specialized assets or characteristics of specific firms").

use of tying or bundling to preserve or extend monopoly power.[260] At the same time, the presence of network effects also may suggest that $E(H_2)$ is low, because the IP owner's return on investment in creating a winning standard far outweighs the cost of development; in appropriately delineated cases, both the frequency and magnitude of false positive errors may be minimal. Cases such as *Sega* may fall within this framework, to the extent that exclusive control over the interface at issue there might have enabled the plaintiff to control follow-up innovation without plausible corresponding benefits to consumers.[261]

In summary, there may be instances in which an IP court should excuse the IP defendant for the purpose of promoting competition, on the basis of potential harms that would not be relevant in an antitrust context. And there may be other, more general, reasons for antitrust courts to be more cautious in their assessment of the relevant costs

---

[260] *See supra* note __.

[261] *See infra* notes __ and accompanying text; *see also* Julie E. Cohen, *Reverse Engineering and the Rise of Electronic Vigilantism: Intellectual Property Implications of "Lock-Out" Programs*, 68 S. CAL. L. REV. 1091, 1192-94 (1995); (arguing that misuse doctrine may be better tailored than antitrust to account for harm to innovation); Note, *Is the Patent Misuse Doctrine Obsolete?*, 110 HARV. L. REV. 1922, 1935-39 (1997); Ramsey Hanna, Note, *Misusing Antitrust: The Search for Functional Misuse Standards*, 46 STAN. L. REV. 401, 418-19 (1994).

Note also that courts sometimes apply the doctrines I refer to as "procompetitive" for reasons other than to promote economic competition—for example, to promote free speech interests. I have no quarrel with this practice, in cases in which free speech or other important noneconomic interests are at stake. From the standpoint of the First Amendment, promoting diversity of expression may be valued for its own sake, even when there are good economic substitutes for the speech in question. *See, e.g.*, Neil Weinstock Netanel, *Copyright and 'Market Power' in the Marketplace of Ideas*, *in* ANTITRUST, PATENTS AND COPYRIGHT: EU AND US PERSPECTIVES __ (Howard Shelanski & François Lévêque eds. 2005) (arguing that "the value of expressive diversity, in the sense of diverse political and artistic viewpoint as opposed to product differentiation *per se*, lies largely in our fundamental, extra-economic commitments to individual self-expression and a robust system of free speech"); *see also* JuNelle Harris, *Beyond Fair Use: Expanding Copyright Misuse to Protect Digital Free Speech*, 13 TEX. INTELL. PROP. L.J. 83 (2004) (arguing for the application of copyright misuse doctrine to promote free speech interests); Kathryn Judge, Note, *Rethinking Copyright Misuse*, 57 STAN. L. REV. 901 (2004) (similar); Note, *supra* note 14 (similar). Excusing the defendant from liability therefore might advance competition in the broader sense of competition in the market of ideas, even though it has no impact on competition in any sense that is cognizable for antitrust purposes.

and benefits of intervention, as compared with IP courts. A false positive in an antitrust case means that the defendant has been wrongly adjudged guilty of violating a federal law which, in theory, can carry criminal penalties. A false positive in the IP setting means only that the IP plaintiff has been wrongly denied the vindication of its rights. To the extent that the antitrust false positive carries with it greater stigma, greater caution may be warranted in the antitrust arena. Moreover, as David McGowan notes, antitrust is concerned with the competitive process, whereas IP law is more concerned with outcomes (i.e., producing the "right" amount of innovation).[262] A generally more aggressive approach to anticompetitive harm, focusing on specific outcomes of the competitive process, therefore might make sense in the IP context. Nevertheless, it remains to be seen how closely the theoretical observations above can be applied to real-world doctrines. The following Part addresses this issue.

V. Implications for IP's Procompetitive Doctrines

The analysis above suggests that there *can* be instances in which the expected benefits of invoking a procompetitive doctrine to excuse liability outweigh the expected costs, even though those costs would be too speculative to form the basis of an antitrust claim. In this Part, I apply the model developed above to the doctrines discussed in Part III. My overarching claim is that, notwithstanding the theoretical observations presented above, the number of cases in which courts should apply procompetitive doctrine to protect "submarkets," aggregate markets, or innovation markets, in a more aggressive manner than antitrust would contemplate, is probably small. More frequently, antitrust-like standards *should* govern when litigants ask courts to vindicate competitive need in IP

[262] *See* McGowan, *supra* note __, at 777.

cases. To the extent that courts sometimes should excuse IP defendants from liability so as to promote competitive ends, however, the application of either fair use or a more narrowly-crafted misuse doctrine would be preferable to the misuse doctrine as it currently exists. I also argue that the Supreme Court's non-antitrust-grounded gloss on trademark law's functionality doctrine probably *does* promote social welfare; but that an even better response might be to eliminate trademark protection for trade dress altogether.

A. Procompetitive Tools in Patent and Copyright

In this section, I take up the issue of whether courts might be justified in applying misuse, fair use, or merger so as to promote competition, even in the absence of demonstrable market harm. I begin by arguing that misuse doctrine, as it currently exists, gives rise to two problems that render most applications of the doctrine inherently suspect. Notwithstanding these problems, I then consider two situations in which misuse or other doctrines nevertheless might promote social welfare by enhancing competition in ways that depart from antitrust standards: first, by overriding IP-owner-imposed restrictions on reverse engineering; and second, by frustrating efforts on the part of IP owners to engage in price discrimination by, for example, attempting to override the first-sale doctrine. I contend that, in some but not all instances, IP law should negate restrictions on reverse-engineering, even without proof of antitrust harm; that the theoretical framework described above does *not* typically provide a sufficiently strong case for restricting efforts on the part of IP owners to price-discriminate; and that, in cases in which excusing the defendant from liability is warranted, fair use (if available) would be preferable to misuse as a doctrinal tool, unless courts incorporate a standing

limitation into the latter. I conclude with some observations on whether it would be useful to incorporate further investigation of market harms into the application of fair use, misuse, and merger as procompetitive tools.

1. Two problems with misuse

A first problem with applying the misuse doctrine against conduct that does not constitute an antitrust violation is that the potential harms from a false positive will often be just as great as, or even greater than, the analogous harms would be in an antitrust setting.[263] As we have seen, allegations of misuse are almost always based upon contractual provisions of one sort or another. But even those provisions that are characterized as misuse per se, such as terms requiring the payment of post-expiration royalties and (some forms of) tying arrangements, typically exhibit countervailing static efficiencies; much more so the practices that fall within the misuse doctrine's rule of reason, such as grantbacks, pools, and field-of-use restrictions. Using the terminology developed in the preceding Part, the loss of these potential benefits comprises a major part of H_2. The other part of H_2, also important, consists of harm to the patent or copyright incentive structure. Unless the decisionmaker can be confident that both parts of H_2 are small, condemning the conduct at issue as misuse will be difficult to justify. As for H_1, if the deadweight loss component is substantial enough, antitrust concerns may arise without the need to consider misuse. But often there will be little reason to expect the other possible components of H_1 to be sufficiently large to justify resort to the misuse doctrine, in a case in which antitrust would not apply. In the typical patent misuse case,

[263] In cases in which the alleged misuse also *would* constitute an antitrust violation, the application of the doctrine is less problematic, but also arguably unnecessary given the viability of antitrust enforcement; and even in these cases there is a risk that the misuse doctrine will result in overdeterrence, conflicting judgments, or other problems, absent a standing limitation. *See infra* notes __ and accompanying text.

for example, there is unlikely to be any significant noneconomic, free-speech value at stake;[264] and while a ruling for the plaintiff might cause potential harm to a "submarket," an aggregate market, or an innovation market, the presence of countervailing H_2 harm should counsel in favor of at least as much caution as would be exercised in the antitrust context.

A second problem is that, because a finding of misuse renders the patent or copyright unenforceable in its entirety, deploying misuse to combat small or speculative competitive harms creates a substantial risk of overdeterrence. Suppose, for example, that a court were to conclude that a contract requiring the user of a copyrighted work not to engage in fair use or reverse engineering constitutes a misuse of the copyright. According to most courts that have applied the misuse doctrine, the copyright would then be unenforceable in its entirety—even as against a defendant whose was use was *not* fair, and even if (as in *Lasercomb*) the defendant was not a party to the agreement embodying the requirement. In all but the most egregious instances, this departure from the standing principles common to both antitrust and the doctrine of unclean hands seems imprudent, insofar as it risks invalidating copyrights without sufficient input from the parties directly affected by the restriction at issue. It also results in the imposition of a sanction on the copyright owner that, in some cases, will bear no relationship to the magnitude of the harm caused by the restriction.[265] A more reasonable interpretation of misuse in both the

[264] *But see* Dan L. Burk, *Patenting Speech*, 79 TEX. L. REV. 99, 150-60 (2000) (arguing that patent law must accommodate free speech values, at least to the extent that software patents read on digitized expression).

[265] Perversely, the lack of a nexus between harm and penalty might sometimes have the opposite effect of underdeterring anticompetitive conduct. Attempts to limit fair use or use of public-domain materials appear to be widespread; it's not difficult, for example, to find works bearing notices that purport to prohibit their reproduction even in part. *See* Patry & Posner, *supra* note __, at 1654-57. Some of these attempts to restrict users may result in anticompetitive harm, but courts may be reluctant to characterize the

patent and copyright contexts would limit its application to matters that relate to the acts or transactions at issue in the litigation, in much the way that unclean hands is limited within the law of restitution.[266] The other alternative would be to discard misuse altogether and to rely instead on a properly reformulated fair use or preemption doctrine, or simply to hold offending contractual provisions unenforceable as a matter of public policy. The potential collateral damage from the application of the misuse doctrine as currently formulated, however, is simply too great, if a narrower approach is available.

2. Restrictions on reverse engineering

Notwithstanding my reservations about misuse doctrine generally, it remains conceivable that there are situations in which misuse or other doctrines might play a role in promoting competition in ways that would elude antitrust. A leading candidate for such a case would be one in which the IP owner attempts to prevent others from engaging in reverse-engineering, either by claiming that the act of reverse engineering constitutes patent or copyright infringement, or by selling or licensing products incorporating the owner's IP upon condition that the purchaser or licensee agree not to reverse engineer the product.[267] A large scholarly literature argues that at least some forms of reverse

---

restrictions as misuse if the penalty is unenforceability of the copyright *tout court*. *See also* Lemley, *supra* note __, at __.

[266] Several commentators have made this observation. *See, e.g.*, Cohen & Lemley, *supra* note __, at 32; Kobak, *supra* note __, at 37-38; Douglas Lichtman, *Property Rights in Emerging Platform Technologies*, 29 J.LEGAL STUD. 615, 636-37 (2000); *cf.* Judge, *supra* note __, at 950-51 (suggesting that courts can avoid overdeterrence by permitting the copyright plaintiff to recover damages but not injunctive relief).

[267] So far, questions concerning the legality of reverse-engineering a patented invention have not arisen much in practice, for several reasons. *See* Cohen & Lemley, *supra* note __, at 21-25; O'Rourke, *supra* note __, at 1227; Robinson, *supra* note __, at 1510 n.219. But patent owners may expressly restrict certain uses, in which case (as in *Mallinckrodt*) the question arises whether the restriction trumps the first-sale doctrine. As far as copyright is concerned, the question may arise, as in *Sega* and *Sony*, whether intermediate copying of software for the purpose of reverse-engineering constitutes infringement or fair use; or, as in *WireDATA*, whether restrictions on data extraction constitute misuse. Finally, the Digital Millennium Copyright Act (DMCA) permits the unauthorized circumvention of technological measures

engineering should be lawful, regardless of the IP owner's objections. Pamela Samuelson and Suzanne Scotchmer, for example, argue that first-mover advantages and the cost of reverse engineering usually provide the IP owner a sufficient opportunity to recoup investment; and therefore that a rule permitting reverse engineering provides a smaller, but still adequate, incentive to invent and disclose.[268] A right to reverse engineer is particularly likely to result in a greater social surplus when network effects are present, again because the rewards to the IP owner are likely to be adequate and the social cost of permitting the IP owner to impede follow-up innovation substantial.[269] A rule forbidding reverse engineering, by contrast, might increase deadweight loss and the cost of follow-up innovation.[270] On balance, they argue, permitting reverse engineering is probably, though not conclusively, efficient in most instances, and courts should be skeptical of efforts to override the right, especially in non-negotiated contracts.[271]

That said, many scholars who support a right to reverse engineer in *some* cases have expressed reservations over interpreting this right too broadly. One possible drawback is that applying a right to reverse engineer with respect to a patented or

---

that restrict access to or use of copyrighted works for some but not all reverse engineering purposes. *See* 17 U.S.C. § 1201(f). Efforts to expand the scope of the IP owner's rights under the DMCA could conceivably give rise to an "anticircumvention misuse" doctrine. *See* Burk, *supra* note __.

[268] *See* Samuelson & Scotchmer, *supra* note __, at 1621.

[269] *See id.* at 1621-26.

[270] *See id.*

[271] *See id.* As stated, however, the effect on incentives depends on reverse-engineering being sufficiently time-consuming and difficult that IP owners have a sufficient opportunity to recoup their costs; the easier reverse-engineering becomes, the more likely that it will negatively impact incentives. *See id.* at 1585-89. Moreover, permitting reverse-engineering may encourage duplicative efforts to reverse-engineer, or efforts on the part of IP owners to make it more difficult to reverse-engineer their products—either of which responses threatens to reduce the social benefits of reverse engineering. *See id.* at 1625. For other discussions favorable to reverse-engineering, see, e.g., Cohen & Lemley, *supra* note __, at 18-37; Brett Frischmann & Dan Moylan, *The Evolving Common Law Doctrine of Copyright Misuse: A Unified Theory and Its Application to Software*, 15 BERKELEY TECH. L.J. 865, 928-30 (2000).

copyrighted work before that work becomes dominant may impede socially beneficial incentives, either by encouraging cloning[272] or by discouraging efficient coordination of follow-up applications.[273] A related point is that, unless legal or contractual restrictions on reverse engineering are widespread within a given industry, there may be little anticompetitive effect from the occasional restriction (although a focus on aggregate impact would be consistent with the analysis provided above). Second, some caution that while a right to reverse engineer for the purpose of creating a product that is complementary to an existing platform may be beneficial, both to society generally (particularly if the product to be engineered has a dominant market position) and to the IP owner specifically (because complements may exacerbate network effects),[274] a right to reverse engineer for the purpose of creating a product that competes with an existing platform is less sound. In the latter case, reverse engineering may enhance social welfare, particularly if the product to be reverse engineered has already achieved market dominance, but the specific facts will be important; a right to reverse engineer in all cases may have a negative impact upon the decision to invest in creating the initial product.[275] On this logic, treating *Sega* (which involved vertical compatibility) and *Sony* (which involved horizontal compatibility) identically may have been a mistake.[276]

---

[272] *See* Weiser, *supra* note __, at 567, 591; *see also* O'Rourke, *supra* note __, at 510 (arguing that a right to reverse-engineer software invites abuse, insofar as the reverse engineer will obtain copyrightable expression along with uncopyrightable data).

[273] *See* Lichtman, *supra* note ___, at 634; *but see* Michael J. Meurer, *Copyright Law and Price Discrimination*, 23 CARDOZO L. REV. 55, 126 (2001) (expressing skepticism over this argument); Samuelson & Scotchmer, *supra* note __, at 1624 n.225 (similar).

[274] *See* O'Rourke, *supra* note __, at 514; Weiser, *supra* note __, at 565-66.

[275] *See* O'Rourke, *supra* note __, at 510; Weiser, *supra* note __, at 567.

[276] *See* Weiser, *supra* note __, at 601-02; *but see* Samuelson & Scotchmer, *supra* note __, at 1622 n.220 (arguing that reverse engineering for horizontal competition is much less common).

But even *Sega* and *Sony* do not create a per se right to reverse engineer software, as noted above. In both cases, the court was careful to articulate that the right would be implicated only if the defendant reverse engineered for a legitimate purpose (presumably, making an identical product would be illegitimate) and if the information could not reasonably have been obtained any other way.[277] Cabining the "right" to reverse engineer even in this modest way, however, renders cases such as *Lasercomb* doubtful. Recall that the agreement in *Lasercomb* forbade licensees from developing computer-assisted die-making software--presumably, whether they engaged in reverse engineering or not.[278] Although a ruling permitting licensees to make compatible products, as in *Sega*, might have made sense had that issue been presented, condemning the license altogether (and rendering Lasercomb's copyright unenforceable) absent further analysis of its anticompetitive effects seems overbroad. To be sure, *anticompetitive effects may have been present*. The restraints at issue had a very long duration and may well have suppressed competition in the market for competing technologies in a way that made consumers worse off. In the antitrust context, however, evidence of anticompetitive effects, with reference to such factors as Lasercomb's market share and the ease of entry into that market, would have been necessary before reaching that conclusion.[279] And while there may have been potential harm to innovation markets or to submarkets, the facts presented in the opinion do not disclose any basis for comparing these harms to the

---

[277] *See supra* text accompanying note __.

[278] *See Lasercomb*, 911 F.2d at 973.

[279] *See* IP GUIDELINES, *supra* note __, § 5.4; 1 HOVENKAMP ET AL., *supra* note __, § 21.7; *see also* James A. Kobak, *A Sensible Doctrine of Misuse for Intellectual Property Cases*, 2 ALB. L.J. SCI. & TECH. 1, 5 (1992) (arguing the license had a "legitimate purpose . . . to prevent . . . surreptitious creation of disguised copies," though "undoubtedly a narrower, less overreaching provision could have served this purpose").

potential harm, to the copyright incentive scheme, of rendering the copyright unenforceable. Lasercomb may well have been overreaching, but absent proof of harm to a well-defined market the application of misuse doctrine in that case, like many others, arguably degenerates into a dubious exercise in "antitrust lite."

These nuances suggest that a broad rule that reverse engineering is always a fair use, or that contractual restrictions on reverse engineering are always a misuse (or otherwise unenforceable) would sweep too far.[280] They also suggest that fair use may have advantages over other doctrinal tools, insofar as the fair use inquiry is inherently more flexible than misuse (with its tendency toward reflexive per se condemnation, as in *Lasercomb* and some of the patent cases) or preemption (which does not easily lend itself to the balancing of costs and incentives).[281] Alternatively, a reformed misuse doctrine that incorporated a standing inquiry into the analysis might be equally capable of making nuanced decisions; it also would be easier to fit within a traditional patent law framework, given that patent, unlike copyright, has never developed a fair use doctrine.

3. Restrictions on price discrimination

[280] To the extent, however, that a court characterizes a particular type of reverse engineering as a fair use of a copyrighted work, it should not permit the IP owner to condition the use of the work upon the user's agreement not to engage in reverse engineering, as in *Bowers v. Baystate Technologies, Inc.*, 320 F.3d 1317 (Fed. Cir.), *cert. denied*, 539 U.S. 928 (2003). To do so enables the parties to frustrate the perceived procompetitive benefits, a type of positive externality, of this application of the fair use doctrine.

[281] *See* Mark A. Lemley, *Beyond Preemption: The Law and Policy of Intellectual Property Licensing*, 87 CAL. L. REV. 111, 157-58 (1999) (noting the difficulty of applying preemption doctrine so as to take into account the incentive/access tradeoff). This is not the place to delve deeply into the law of preemption. Suffice to say that federal IP law may preempt state laws, including the application of state contract law, that would negate important aspects of federal law; but the legal standards are governed by many technicalities and do not easily lend themselves to case-by-case consideration of factors such as the magnitude of H_2.

Another difficult problem is presented in cases such as *Mallinckrodt, Inc. v. Medipart, Inc.*,[282] *ProCD, Inc. v. Zeidenberg*,[283] and *Lexmark International, Inc. v. Static Control Components, Inc.*[284] Of these, only *Mallinckrodt* was litigated as a misuse case—*ProCD* was decided on preemption grounds, and *Lexmark* on the basis of the Digital Millennium Copyright Act (DMCA)—but each case involved a variation on a common pattern. In *Mallinckrodt*, the plaintiff forbade purchasers of its patented devices from reusing those devices at all;[285] in *ProCD*, the plaintiff forbade purchasers of the "consumer package" version of its uncopyrightable database from using the product for a commercial purpose;[286] and in *Lexmark*, the plaintiff sold toner cartridges that were engineered to disable their use in the event they were refilled by anyone other than Lexmark.[287] In each case, the plaintiff was attempting to engage in price discrimination. "Heavy" users of the device at issue in *Mallinckrodt*, the database at issue in *ProCD*, or the Lexmark printers presumably would value the product more, and would be willing to pay a higher price for each use of it, than would "light" users. For the IP owner to exploit this potential profit opportunity is not easy, however. Offering the product to different users at different prices is not feasible if heavy users can pass themselves off as light users, or if light users can buy at a low price and then resell to high users at a small mark-

[282] 976 F.2d 700 (Fed. Cir. 1992).

[283] 86 F.3d 1447 (7th Cir. 1996).

[284] 387 F.3d 522 (6th Cir. 2004).

[285] *See* 976 F.2d at 702.

[286] *See* 86 F.3d at 1450.

[287] *See* 387 F.3d at 529-30. Lexmark also sold a refillable cartridge, at a higher price. *See id.* at 530.

up.[288] Alternatively, offering the product on condition that the user pay a periodic fee based on use gives rise to monitoring problems. In *Mallinckrodt* and *Lexmark*, the plaintiffs sought to avoid these problems by requiring purchasers to buy a new product for each use; in *ProCD*, the plaintiff instead charged different prices to commercial users and noncommercial users, and forbade the latter from engaging in arbitrage.

The resolution of these cases has been inconsistent. As noted above, the Federal Circuit held that Mallinckrodt's single-use policy did not amount to patent misuse, unless the defendant could establish that the policy violated the rule of reason.[289] To the same effect, albeit under a different doctrinal guise, the Seventh Circuit held that ProCD's policy was enforceable as a matter of state contract law and was not preempted by federal copyright law.[290] By contrast, in *Lexmark* the Sixth Circuit held that the DMCA did not forbid the defendant from circumventing the single-use design feature.[291] Antitrust law, however, probably would not have condemned any of the practices at issue, absent proof of market power and anticompetitive effect.[292] For advocates of price discrimination,

---

[288] *See ProCD*, 86 F.3d at 1450.

[289] *See Mallinckrodt*, 962 F.2d at 708-09.

[290] *See ProCD*, 86 F.3d at 1450-55.

[291] *See Lexmark*, 387 F.3d at 545-51; *see also* Chamberlain Group, Inc. v. Skylink Technologies, Inc., 381 F.3d 1178, 1192-1204 (Fed. Cir. 2004) (holding that the defendant did not violate the DMCA by marketing a universal remote control to purchasers of plaintiff's garage door openers, even though the remotes enable users to circumvent technology that controls access to code embedded in the openers), *cert. denied*, 125 S. Ct. 1669 (2005). In a related copyright context, the Ninth Circuit in *MAI Sys. Corp. v. Peak Computer, Inc.*, 991 F.2d 511, 516-19 (9th Cir. 1993) held that merely running a lawfully-made copy of computer software without permission of the copyright owner creates an infringing copy of the software. This ruling temporarily allowed some owners of copyrighted software to control aftermarkets for computer repairs, insofar as repair would entail turning the computer on to see how the system works; but another portion of the DMCA itself overruled this particular result as it relates to maintenance and repair. *See* DMCA, Pub. L. No. 105-304, 112 Stat. 2879 (1998) (amending 17 U.S.C. § 117 to overrule in part *MAI Sys. Corp.*).

[292] Nor would the Robinson-Patman Act likely apply, despite its prohibition on price discrimination. *See* 15 U.S.C. § 13 (2000). That prohibition applies only under narrow circumstances, and it is not usually

antitrust law's reluctance to intervene is just as well. Absent the pricing strategies employed in these cases, the plaintiffs would have sold their products at a unitary price higher than the price offered to "light" users, with the net effect that light users would have been worse off.[293] Antitrust sensibly avoids the false positives in such cases and, one might argue, IP law would be wise to follow suit.

There are other considerations to take into account, however, before reaching a firm conclusion. First, although first-degree price discrimination, in which a seller with market power offers each individual buyer the maximum price she is willing to pay above marginal cost, is economically efficient in comparison with offering a single price above marginal cost,[294] first-degree discrimination is almost impossible to implement[295] and was not at issue in any of the above cases. Instead, each case involved second-degree discrimination, in which different versions of a product were offered to consumers at different prices.[296] The welfare effects of this latter type of price discrimination, however, are ambiguous. From the standpoint of static efficiency, imperfect price discrimination may either increase or decrease output, and hence social welfare, depending upon the shape of the relevant demand and cost curves.[297] The effects on

---

viewed as being consistent with the general thrust of the antitrust laws towards the goal of economic efficiency. *See* HOVENKAMP, *supra* note __, §§ 14.6a1, 14.6d; Meurer, *supra* note __, at 1882-83.

[293] In *Mallinckrodt*, this would mean selling the device at a single price on the expectation that it would be used *x* times before being disposed of.

[294] *See* TIROLE, *supra* note __, at 1234-37.

[295] *See* Meurer, *supra* note ___, at 59 (noting that, for price discrimination to occur, the seller must enjoy market power and be able to accurately estimate buyers' reservation prices, and arbitrage must be difficult).

[296] *See id.* at 69-71. Third-degree price discrimination exists when a supplier offers a product to different groups at different prices, based on a prediction of different groups' willingness to pay. *See id.* at 71-75.

[297] *See* LANDES & POSNER, *supra* note ___, at 40, 378, 389; Meurer, *supra* note ___, at 73-75. In addition, for any form of price discrimination to work, the supplier must incur (or shift) the cost of gauging

dynamic efficiency are similarly ambiguous. Because price discrimination increases the portion of consumer surplus accruing to the monopolist, it encourages efforts to obtain monopolies, including IP monopolies. To the extent that investment in the production of some IP goods, such as drugs, would be inadequate absent the prospect of obtaining monopoly rights, this encouragement is beneficial. On the other hand, to the extent the race to acquire monopoly position encourages overinvestment and duplicative research, it is socially wasteful.[298] Second, even in cases in which price discrimination increases output over what it otherwise would be, the effect on dissemination of the goods produced may be ambiguous. In particular, when price discrimination strategies are permitted to override the first-sale doctrine, consumers who otherwise would have obtained the product at *below* marginal cost (through sharing, resale, or gift) may lose access to the product.[299] For example, assume that in the absence of price discrimination I would purchase a product at the monopoly price of $10 and then resell to someone else for a price below the product's marginal cost. Under a regime of price discrimination coupled with a restriction on resale or reuse, I might pay less to obtain the product but the hypothetical second owner may not obtain the product at all, if she is unwilling to pay a price at least equal to its marginal cost. This loss to the second owner may be irrelevant as long as economic efficiency is defined in terms of willingness-to-pay.[300] To the extent

---

consumer preferences and enforcing restrictions against arbitrage. Presumably a supplier who engages in price discrimination expects to better off, even in the presence of these costs, than it otherwise would be, but the social welfare calculus must take these costs into account as well. *See* LANDES & POSNER, *supra* note __, at 389; Meurer, *supra* note __, at 101.

[298] *See* LANDES & POSNER, *supra* note __, at 378, 389; TIROLE, *supra* note ___, at 139, 149; Meurer, *supra* note ___, at 100-02.

[299] *See* Thomas F. Cotter, *Memes and Copyright*, 80 TULANE L. REV. __, __ (forthcoming 2005).

[300] *See id.*

that the second owner's willingness to pay is constrained by her ability to pay, however, or her use of the product would confer unquantifiable positive externalities (e.g., spillover educational benefits) upon third parties, a more expansive social welfare criterion might prefer the first-sale regime to the price discrimination regime.[301] Preventing price discrimination for these latter purposes, of course, would have little to do with competition policy as such but rather would be rooted in other values.

Notwithstanding these ambiguities, or perhaps because of them, I remain skeptical of the view that IP law should inhibit price discrimination schemes as a matter of course. To be sure, there may be some cases in which doing so would enhance social welfare. For example, Michael Meurer argues that ProCD's price discrimination policy probably was welfare-reducing, based upon his assessment that (1) little incentive is necessary to induce production of the type of database at issue in that case, and (2) duplicative efforts to produce similar databases promise little social value in return for the cost.[302] As Meurer concedes, however, there is at present little empirical analysis supporting or negating either proposition.[303] Advocates of rights in databases clearly disagree with Meurer on the incentive issue, while others may doubt the ability of courts to make the correct assessment of costs and benefits without hindsight bias.[304] A further

---

[301] *See id.*

[302] *See* Meurer, *supra* note __, at 105-08.

[303] *See id.* at 108.

[304] Doctrinally, it is also difficult to account for the incentive effect when the issue is, as in *ProCD*, one of preemption rather than misuse. *See supra* note __. Note also that in *ProCD* and other cases involving so-called shrinkwrap licenses, if the license term at issue is unenforceable as a matter of contract law, because it lacks mutual consent, there is nothing for copyright law to preempt. The Seventh Circuit in *ProCD* held that the license at issue there was enforceable as a matter of contract law, *see* 86 F.3d at 1450-53, but the opinions of other courts and commentators may differ.

consideration is that the invalidation of contractual restrictions on use or resale may induce producers to incorporate technological use restrictions into their products.[305] The substitution of technology for contract enables producers to evade the legal problems arising from contractual restrictions, unless courts decide to monitor design choices—a task for which they are probably not well-positioned.[306] Thus, despite suggestions that misuse doctrine plays a role in filling gaps left by antitrust,[307] and despite the fact that IP law in theory *could* fill a gap by condemning *some* inefficient price discrimination-enhancing contracts that would evade antitrust scrutiny, the preceding analysis leads me to doubt that misuse (or preemption) would often fare any better than antitrust in performing this function.[308]

4. What Is to Be Done?

[305] *See* Lichtman, *supra* note __, at __; Meurer, *supra* note __, at 107; Robinson, *supra* note __, at 1513.

[306] *See* Robinson, *supra* note __, at 1519. Precisely out of concern that courts become entangled in questions concerning design choices, the theory of predatory innovation—i.e., that innovation may violate the antitrust laws, by raising rivals' costs without providing a commensurate social benefit—has not won widespread favor, despite its potential theoretic appeal.

[307] *See* Burk, *supra* note __, at 1122; Frischmann & Moylan, *supra* note __, at 872.

[308] *See also* Robinson, *supra* note __, at 1506-07 (arguing that, in light of the ambiguous effects of price discrimination, construing the first-sale doctrine as a nonmodifiable rule is unwarranted). I should qualify the statement in the text above in two particulars, however. First, restrictions on reverse engineering, which I am somewhat concerned about, can themselves be viewed as means for enabling price discrimination. For example, consider Sega's attempt to prevent Accolade from reverse-engineering the source code for Sega's consoles. Sega sold the consoles, which are useless without games to play on them, with the expectation that it would profit from the sale of Sega-authorized video games to be played on the console. Users' willingness to buy multiple games would be a measure of the intensity of their preference for the Sega system. *See* Meurer, *supra* note __, at 117-18; O'Rourke, *supra* note __, at 533; Hanna, *supra* note __, at 433-34. Thus when measures to prevent reverse engineering are unenforceable, the Segas of the world may respond by raising console prices, or by designing consoles to be yet more difficult to reverse engineer. Both responses potentially have negative welfare consequences. Nevertheless, I continue to find persuasive the arguments that reverse-engineering restrictions in some instances reduce welfare, even if price discrimination generally should be permitted. Second, perhaps in some cases adherence to the first-sale doctrine is justified for reasons related to free speech or other constitutional values. *See infra* text accompanying notes __.

The preceding analysis suggests that there may be cases in which the balance of expected costs and benefits is such that a court would enhance social welfare by excusing the patent or copyright defendant from liability, even though the potential anticompetitive harm would not be cognizable, or would be too speculative, in the analogous antitrust context. Even so, the analysis sounds a skeptical note that there are a great many such cases, principally because courts often cannot be sufficiently confident that the potential harm from ruling for the defendant will be minimal. A related question therefore is how extensive a court's scrutiny of the relative harms should be. The more intensive the factual inquiry, the more confident one may be of the outcome. On the other hand, even with a very extensive examination of the facts, it may be impossible to quantify the relevant harms--particularly the harms to the IP incentive structure. Ideally, the system would encourage investment in further development of the facts, up to the point at which the returns on the investment promise a positive payoff—that is, up to the point at which one would expect another increment of investigatory cost to produce a commensurate benefit, in terms of reducing the risk of false positives or false negatives. But it is unlikely that courts can make more than a rough guess of the potential benefits from further investigation. Clearly, if it would cost some nominal amount to reduce the variance of the expected cost/benefit tradeoff to zero, it would be worth doing so. Beyond this truism, however, lurks the unhappy reality that the variance is likely to be unknown and unknowable, as is the effect of spending another increment of cost in reducing it. A few observations nevertheless may be relevant.

The first is that the resolution of many of the preceding questions may depend upon which party has the burden of proof. Under current law, misuse, fair use, and

merger are usually viewed as affirmative defenses. Given this starting point, in an ambiguous case—that is, one in which the magnitude of the relevant harms, including the harm to the incentive structure, is an open question—the IP plaintiff should prevail. To be sure, this proposal merely begs the question of where the burden of proof *should* lie. Allocating the burden to the IP defendant, as under current law, is consistent with the assumption that H_2 is in general high. Allocating the burden to the IP plaintiff to *disprove* misuse or fair use, on the other hand, would raise the cost of enforcing IP rights and thus reduce the value of those rights; at the same time, it would reduce the incidence of false negatives and increase the incidence of false positives. In part the resolution of this dilemma must depend upon how important one views the incentive structure to be.[309]

Second, there may be several situations in which it clearly does not make sense to invest heavily in additional investigation of the pro- or anticompetitive effects of excusing the IP defendant from liability. One would be a case in which the stakes are simply too low to justify the additional cost. For example, suppose that preliminary evidence suggests that $E(H_1)$ is \$10,000, $E(H_2)$ is \$5,000, and that the expected cost of an additional increment of litigation relating to proof of market harm is \$50,000. Even if these litigation costs could reduce uncertainty to zero, it seems unlikely that the additional cost is worthwhile. (If the cost of additional litigation is not discontinuous, as in this example, some additional cost may be justified, though as noted above it may be impossible to estimate with any accuracy just how much.) In this type of case, if the

[309] Alternatively, policymakers could explore burden-shifting mechanisms as a means for better vindicating noneconomic interests such as free speech. A few scholars have recommended such an approach in some limited fair use contexts. *See, e.g.*, Kenneth D. Crews, *Fair Use of Unpublished Works: Burdens of Proof and the Integrity of Copyright*, 31 ARIZ. ST. L.J. 1, 68 (1999) (arguing that, when the defendant has copied an unpublished manuscript, the burden of rebutting fair use should rest upon the plaintiff, in order to vindicate free speech interests). I plan to take this matter up in greater depth in a separate paper.

court is sufficiently confident that $E(H_1) > E(H_2)$, without further investigation, it should excuse the IP defendant from liability; if not, and if the burden of proof rests on the defendant, the court should rule for the IP plaintiff.

Another type of case in which further investigation of market harms may not be justified is one in which the court can be sufficiently confident without further evidence that (1) $E(H_2)$ is close to zero, and (2) $E(H_1)$ is greater than zero. In such a case, even if there is considerable variance in the estimate of $E(H_1)$, the cost of reducing that variance will not produce any social benefit. Courts should not boldly assume that these cases are common, however; to do so invites judges to substitute their intuition for the judgment of Congress that IP rights encourage a variety of social goods. That said, there are nonetheless several plausible instances in which economic theory suggests that the effect upon incentives of excusing the IP defendant from liability should be minimal.[310] When that is clearly the case, social welfare will increase even if the H_1 harm to be avoided is somewhat speculative. *Sega* and *Lotus v. Borland*, conceivably, are two such cases.[311] So too are cases applying the merger doctrine to exclude from the scope of copyright protection code that is necessary for interoperability; works for which other incentives, such as direct government largesse, are likely to be sufficient; and works involving minimal creativity, such as sweepstakes instructions.[312] Even in cases such as these, however, unless the stakes are sufficiently low it may be unwise to preclude the parties from offering evidence relevant to the plaintiff's market dominance, the existence of

[310] *See supra* text accompanying notes __.

[311] *See supra* notes __ and accompanying text.

[312] *See supra* notes __ and accompanying text.

network effects, and other evidence bearing on the magnitude of the expected harms. Courts should not short-circuit the inquiry by taking judicial notice that there is no likely impact upon incentives, if in fact evidence might reasonably cast doubt on this intuition. On the other hand, elaborate analysis of market definition may not always be the appropriate focal point. In an appropriate case, analysis may properly center upon the incentive question directly, e.g., are there sufficient non-copyright incentives in place to induce production of similar works, and are there adequate alternative means of accessing the work without copying.[313]

In most other instances, however, which raise difficult questions of misuse or fair use—particularly cases involving restrictions on reverse engineering or price discrimination—courts probably should require some proof of anticompetitive effects before excusing the IP defendant from liability. To be sure, this will raise the cost of litigating such cases, but the only way to seriously engage the issue of whether there is a plausible procompetitive rationale for excusing the IP defendant from liability is by predicting the market consequences using the available tools of economics. Even so, it is conceivable that some sort of sliding scale might be appropriate. A relatively small prospect of H_2 harm might enable courts to be more aggressive in permitting unauthorized uses so as to minimize harm to submarkets, aggregate markets, or innovation markets. Resort to a full-blown antitrust inquiry may be unnecessary in some

[313] Yet another case in which elaborate investigation of market harms would be unjustified would be one in which the restriction at issue is deemed problematic for non-competition policy related reasons, such as free speech. *See supra* note __. Just because H_1 may comprise a substantial free-speech component, however, does not necessarily mean that the resolution of the matter will be easy. If the dynamic efficiency component of H_2 is also large, for example, a ruling for the IP defendant might actually decrease the quantity of speech in the long run. See Harper & Row Publishers, Inc. v. Nation Enters., 471 U.S. 539, ___ (1985). There may be no entirely satisfactory way of resolving the tradeoff in such a case.

cases, but the case law should be allowed to develop with the understanding that the cost of forgoing such an inquiry may be a greater incidence of false positives.

B. Functionality

In contrast to much of the preceding analysis, comparison of the costs and benefits relevant to trademark law's functionality doctrine is relatively straightforward. First off, there are many reasons to think that, in the context of product design trade dress specifically, $E(H_2)$ is often small. As noted, the rationales for protecting trademarks of any kind is that consumers use trademarks as source-identifiers, and thus are likely be harmed by the unauthorized, confusing use of similar marks; thus, in the absence of trademark protection, firms would have little incentive to invest in maintaining quality control. As applied to product design trade dress, however, these rationales are relatively weak. For one thing, the standard of liability for infringement of any type is *likelihood* of confusion, not actual confusion; and for a use to infringe, it need only be likely to confuse a "substantial portion," not a majority, of the relevant class of consumers.[314] Survey evidence purporting to show that a substantial minority of that class is likely to be confused is often sufficient to prove liability.[315] Furthermore, trademark owners are not required to show that the confusion is material to consumers' purchasing decisions.[316] To be sure, all of these features of trademark law apply to conventional trademarks as well, and for good reason. A standard that imposed liability only when the trademark owner proved actual harm to consumers would dramatically increase the cost of

[314] *See* RESTATEMENT (THIRD) OF UNFAIR COMPETITION § 20 cmt. g (1995).

[315] *See id.*; 3 MCCARTHY, *supra* note __, § 23.01[2].

[316] *See* Glynn S. Lunney, Jr., *Trademark Monopolies*, 48 EMORY L.J. 367, 482-83 (1999); *see also* Bone, *supra* note __, at 2147-48 (noting that trademark doctrine does not take into consideration the magnitude of the harm resulting from confusion).

trademark enforcement; at the same time, in most instances it (probably) would confer little benefit, because potential defendants typically have a large number of noninfringing conventional marks (e.g., words) from which to choose. Thus, as Bone notes, even if the liability standard is overinclusive (in the sense of condemning more conduct than would an actual-harm standard), the social cost of overinclusiveness may be minimal, and the benefits (in terms of low enforcement costs) substantial.[317] With respect to product design trade dress, however, the liability standard is even more overinclusive, because consumers frequently encounter trade dress in conjunction with other source identifiers (such as word marks). The potential for harmful confusion is therefore reduced, even though it may not be eliminated.[318] Finally, producers of goods bearing word marks and having a distinctive trade dress would still have a robust incentive to invest in quality control, if other firms could copy their trade dress but not their word marks. Purchasers would still encounter the word marks and could be expected to rely upon them in making future purchasing decisions.[319] On balance, the potential cost of false positives in product design trade dress cases may be relatively low.

On the other hand, in the typical case H_1 may not be very high either. Unless the exercise of exclusive rights to a particular product design enables the rights owner to exercise power over a distinct market, the deadweight loss from protection will be small.

[317] *See* Bone, *supra* note __, at 2136.

[318] Of course, some consumers may ignore the word marks, or they may encounter products in post-sale contexts in which other source identifiers are absent. But there is clearly *less* potential for harmful confusion than in a case involving a deceptively similar word affixed to similar goods. Disclaimers too could reduce, even if they did not eliminate, some residual confusion. RESTATEMENT (THIRD) OF UNFAIR COMPETITION, § 35 cmt. d (1995) (discussing disclaimers).

[319] *See* Bone, *supra* note __, at 2146-47; *but see id.* at 2172 (stating that the risk of post-sale confusion may affect some firms' decisions to invest in goodwill).

Nevertheless, even harm to a "submarket" may be worth vindicating, if the countervailing benefit to consumers and producers is even smaller. And in the aggregate the H_1 harm may not be so small after all. If the relevant market can accommodate only a small number of variations on a particular design, allowing a correspondingly small number of firms to appropriate them may create a meaningful barrier to entry. Even the possibility of being sued for trade dress infringement may create some deterrent to entry, as the Supreme Court suggested in *Wal-Mart*.[320] Finally, there is some risk that, if courts are unwilling to require proof of market definition, they will rely upon intuition to define product markets more broadly than necessary, thus allowing trade dress owners to assert exclusive rights over products that *do* inhabit discrete markets.[321] In such cases, the deadweight loss may be substantial. Once again, though, one can only speculate about the magnitude of H_1 in the absence of market evidence.

The other factor to take into account, however, is the cost of defining the market. If this cost outweighs the benefits of greater certainty, the better choice is not to incur the cost. The preceding analysis suggests that the stakes of the typical product design trade dress case are not sufficient to justify the cost of turning every such case into a miniature antitrust dispute; but if so, then what should the default rule be? Even if courts are reluctant to inject the expense of formal market definition into trademark litigation, they could still require *relatively* strict evidence whether the plaintiff's trade dress inhabits a discrete market—for example, evidence that few alternative product designs exist—before concluding that the trade dress is functional. (Logically, however, this choice

[320] *See* Wal-Mart Stores, Inc. v. Samara Bros., 529 U.S. 205, 213-14 (2000).

[321] *See* Bone, *supra* note __, at 2176-77.

would still require some grounds for determining whether another design *is* a feasible alternative, and thus would avoid an analysis of market harm only by ignoring the question.) Or they could require relatively little proof—for example, evidence that the design feature at issue is necessary to the use of the plaintiff's product, without inquiry into whether adequate market substitutes for that product exist—on the theory that it is better to err on the side of access. To complicate matters further, might there be *some* cases in which the stakes are sufficiently weighty to merit inquiry into market harm, à la antitrust? If so, should courts decline to consider such evidence nonetheless, on the theory that it would open the door to similar evidence in other cases in which the plaintiff might use the threat of substantial litigation costs to keep potential defendants at bay?

Rightly or wrongly, post-*TrafFix* courts are directed to classify product design as functional on the basis of relatively scant evidence of market harm. This standard makes sense if (1) one can be reasonably confident that $E(H_1) > E(H_2)$ in most cases, and (2) the social cost of obtaining greater certainty outweighs the social benefit of dispelling this presumption in the occasional case. If the above analysis of $E(H_1)$ and $E(H_2)$ is correct, these conditions may well be present in the vast majority of product design cases: $E(H_2)$ may be close to zero, $E(H_1)$ somewhat larger even though speculative, and the cost of greater certainty unjustified in light of the small aggregate harm resulting from the (occasional) false positive. A logically compelling further step would be to urge courts to apply narrow product definitions in cases involving questions of aesthetic functionality. For that matter, Bone's surmise that the costs of protecting product design trade dress under *any* circumstances outweigh the minimal benefits may well be correct, though as he recognizes the matter cannot be conclusively proven absent the elusive empirical

evidence.[322] If the above analysis is correct that H_2 is often close to zero, however, trade dress protection arguably imposes more costs than benefits across the board, and social welfare would be greater if protection were eliminated altogether.

VI. Conclusion

Courts sometimes invoke a variety of IP doctrines for the express purpose of promoting competition, but without the careful analysis of anticompetitive consequences that is often a hallmark of antitrust litigation. Nonetheless, I have argued that IP's more flexible approach can make sense in some limited contexts, because IP law can and should take into account a variety of harms that would be entitled to little weight in an antitrust context. More specifically, the expected harm, including error costs, of ruling for the IP defendant will sometimes be sufficiently low as to counsel in favor of excusing the defendant from liability, even though the potential anticompetitive harm resulting from a ruling for the plaintiff would be unduly speculative or not cognizable at all from the perspective of antitrust. These considerations also necessarily affect the analysis of how much courts and litigants should invest in reducing the variance of expected outcomes. Again, in some IP contexts it may not be wise to invest more heavily in obtaining information about the anticompetitive effects of the exercise of IP rights, either because the stakes are too low or because such information would be of limited value to estimating the magnitude of the harm at issue.

[322] As noted above, Bone surmises that courts often intuitively adopt broader product definitions than the economic evidence, if it were to be conducted, would suggest, thus in effect permitting trade dress owners to exert exclusive rights over products that inhabit discrete markets. *See id.* at 2146-47 Unless one actually produces the economic evidence, however, the surmise that the social costs of trade dress protection outweigh the social benefits remains just that—a surmise—as Bone recognizes. *See id.* at 2183-85. In any event, by allocating the burden of proving nonfunctionality to plaintiffs, the law takes a tentative step in favor of resolving ambiguous cases in favor of IP defendants.

Nothing in this analysis, however, suggests that courts should decide cases by indiscriminately imposing their own intuitions concerning the optimal scope of IP rights. In the patent context in particular, courts should be reluctant to take upon themselves the duty of promoting competition by applying a greatly expanded misuse doctrine. With respect to copyrights, however, the balancing of interests permitted by the fair use and merger doctrines may counsel in favor of a more aggressive procompetitive stance, in some cases, than would be advisable in an antitrust context. Similarly, trademark law's functionality doctrine probably serves a useful social function insofar as it weakens protection for product design trade dress, because even a weak potential for causing anticompetitive harm most likely outweighs the social benefit of reducing marginal consumer confusion. Further research may shed light on the question of whether antitrust law itself should consider the unique costs and benefits of IP rights, for example in evaluating unilateral refusals to license IP rights. Developing workable standards for evaluating such claims may not be easy, but if analysts focus on the real interests at stake such efforts over time may bear fruit.