

4-30-2014

The First Sale Doctrine and the Economics of Post-Sale Restraints

Ariel Katz

Follow this and additional works at: <https://digitalcommons.law.byu.edu/lawreview>



Part of the [Intellectual Property Law Commons](#)

Recommended Citation

Ariel Katz, *The First Sale Doctrine and the Economics of Post-Sale Restraints*, 2014 BYU L. Rev. 55 (2014).

Available at: <https://digitalcommons.law.byu.edu/lawreview/vol2014/iss1/4>

This Article is brought to you for free and open access by the Brigham Young University Law Review at BYU Law Digital Commons. It has been accepted for inclusion in BYU Law Review by an authorized editor of BYU Law Digital Commons. For more information, please contact hunterlawlibrary@byu.edu.

The First Sale Doctrine and the Economics of Post-Sale Restraints

*Ariel Katz**

ABSTRACT

The first sale doctrine limits the exclusive rights that survive the initial authorized sale of an item protected by intellectual property (IP) rights, and therefore limits the ability of IP owners to impose post-sale restraints on the distribution or use of items embodying their IP. While the doctrine has deep common law and statutory roots, its exact rationale and scope have never been fully explored and articulated. As a result, the law remains somewhat unsettled, in particular with respect to the ability of IP owners to opt-out of the doctrine and with respect to the applicability of the doctrine to situations of parallel importation. This Article provides answers to these unsettled issues. By applying insights from the economics of post-sale restraints, the Article shows that the main benefits of post-sale restraints involve situations of imperfect vertical integration between coproducing or collaborating firms, which occur during the production and distribution phases or shortly thereafter. In such situations, opting out of the first sale doctrine should be permitted. Beyond such limited circumstances, however, the first sale doctrine promotes important social and economic goals: it promotes efficient long-term use and preservation of goods embodying IP and

* Associate Professor, Innovation Chair–Electronic Commerce, Faculty of Law, University of Toronto. This Article is based on a paper given at the Exhaustion and First Sale in IP Symposium, High Tech Law Institute, Santa Clara Law School, Nov. 5, 2010, at the 2011 Intellectual Property Scholars Conference, and the 2012 Next Generation of Antitrust Scholarship Conference. I wish to thank Eric Goldman, Mark Lemley, Mark McKenna, Anne Layne-Farrar, Mark Patterson, Steven Edwards, and other symposium participants for their valuable comments and insights, as well as to thank Norman Siebrasse, and my colleagues Peggy Radin, Abraham Drassinower, Ed Iacobucci, and Michael Trebilcock for their advice. Special thanks go to Michael Laskey and Nicholas Van Exan for their superb research assistance.

facilitates user-innovation. Therefore, contrary to some other views, I conclude that the economics of post-sale restraints confirm the validity and support the continued vitality of the first sale doctrine.

I. INTRODUCTION.....	57
II. THE FIRST SALE DOCTRINE: DIFFERENT FORMULATIONS.....	61
<i>A. No First Sale Doctrine</i>	61
<i>B. Weak First Sale Doctrine</i>	62
<i>C. Moderate First Sale Doctrine</i>	62
<i>D. Strong First Sale Doctrine</i>	63
<i>E. Strongest First Sale Doctrine</i>	63
III. THE ORIGINS.....	64
<i>A. The Antitrust Legacy of Bobbs-Merrill</i>	65
IV. MODERN ANTITRUST APPROACH TOWARDS POST-SALE RESTRAINTS	70
V. THE FLAW IN THE PARALLEL TRADE DEBATE.....	74
<i>A. Price Discrimination</i>	76
<i>B. Other Efficiencies of Post-Sale Restraints</i>	84
1. <i>Parallel trade and pre-sale and post-sale services</i>	84
2. <i>Parallel trade and its effect on positional goods</i>	85
<i>C. The Irrelevance of Antitrust Insights to the Parallel Trade Debate</i>	86
VI. THE PROPERTY/CONTRACT FLAW	89
<i>A. The Costs and Benefits of Infringement Remedies</i>	94
VII. THE (NON)-COASIAN FLAW.....	100
VIII. JUSTIFYING THE FIRST SALE DOCTRINE	109
<i>A. Preservation and Resource Waste</i>	110
<i>B. Protecting the “Innovation Wetlands”</i>	112

IX. APPLICATION.....	117
A. Costco v. Omega	117
B. Kirtsaeng v. John Wiley & Sons, Inc.	119
C. Vernor v. Autodesk	126
D. UMG Recordings, Inc. v. Augusto.....	130
E. Capitol Records v. ReDigi	133
F. Summary	136
X. CONCLUSION.....	141

I. INTRODUCTION

The first sale doctrine in copyright law, also known as the exhaustion rule of intellectual property (IP) rights, limits the power of IP owners to control the downstream distribution and use of their products or copies of their products that bear their trademark or embody their invention or work. When such rules apply, the patent owner's right to sell a product embodying the invention,¹ the copyright owner's right to distribute copies of her work,² or the trademark owner's right to sell products bearing the trademark³ are terminated after the first authorized sale of the genuine product or work. As a result, any subsequent sale, rental,⁴ or other disposition of ownership or possession (and in the case of patents, use of the patented invention)⁵ does not require the authorization of the owner of the IP right.

Occasionally, intellectual property owners have wished to exert greater control over the downstream distribution or use of their works and products and have looked for ways to work around the

1. Patent Act, 35 U.S.C. §§ 1–376 (2010).

2. Copyright Act, 17 U.S.C. §§ 101–1332 (2010).

3. Lanham Act, 15 U.S.C. §§ 1051–1141 (2010).

4. Software and sound recordings are exceptions to this rule. Copyright Act, 17 U.S.C. § 109; Computer Software Rental Amendments Act of 1990, Pub. L. No. 101–650, 104 Stat 5089 (codified as amended at 17 U.S.C. § 109 (1994)).

5. In addition to making, offering for sale, and selling the patented invention, the Patent Act grants the patentee an exclusive right to “use” the patented invention. 35 U.S.C. § 154 (2010). Copyrights or trademarks do not confer such general exclusive right to “use.”

first sale doctrine using both technological and legal means, or to limit its scope by statutes. So far, these attempts have been only partially successful, resulting in a set of incomplete, and often incoherent and confusing, rules and exceptions. The Copyright Act, for example, contains ambiguous language with regard to the application of the doctrine to copies that were made outside the United States. This ambiguity has required the Supreme Court to hear three cases within the last fifteen years until the Court, in a divided decision, resolved that the doctrine does not discriminate between copies that were made in the United States and those that were made abroad.⁶ The statutory language has also been interpreted as confining the application of the doctrine only to copies that are “owned” rather than “licensed,”⁷ or to render the doctrine inapplicable to the resale of digital copies apart from the resale of the particular medium in which they are embedded.⁸ The language is also silent about the doctrine’s effect on contractual workarounds between owners and users and the effect of such contracts on third-party purchasers. In patent law, even though the statutory language is simpler,⁹ courts still grapple with defining the exact scope of exhaustion, and the Court has refrained from expressing an opinion on the validity of contractual workarounds.¹⁰ With the increasing

6. *Quality King Distribs., Inc. v. Lanza Res. Int’l, Inc.*, 523 U.S. 135, 154 (1998) (Ginsburg, J., concurring) (concurring on the basis that the imported copies were made in the United States); *Costco v. Omega*, 541 F.3d 982 (9th Cir. 2008), *affm’d by an equally divided court*, 131 S. Ct. 565 (2010) (holding that the first sale doctrine applies only to copies made in the United States); *Kirtsaeng v. John Wiley & Sons, Inc.*, 133 S. Ct. 1351, 1357–71 (2013) (holding that the first sale doctrine is not limited to copies lawfully made in the United States and equally applies to copies of a copyrighted work lawfully made abroad).

7. *Vernor v. Autodesk, Inc.*, 621 F.3d 1102 (9th Cir. 2010).

8. *Capitol Records, LLC v. ReDigi Inc.*, 934 F. Supp. 2d 640 (S.D.N.Y. 2013).

9. The Patent Act unambiguously grants an exclusive right to make, use, or sell the invention, and exhaustion rules, developed by the courts, have never been codified. Somewhat paradoxically, the lack of codification saves courts from the need to grapple with ambiguous language such as that of the Copyright Act.

10. *Quanta Computer, Inc. v. LG Elecs., Inc.*, 553 U.S. 617 (2008) (leaving unanswered the question of contracting around the exhaustion doctrine, and noting that “the authorized nature of the sale to Quanta does not necessarily limit LGE’s other contract rights. LGE’s complaint does not include a breach-of-contract claim, and we express no opinion on whether contract damages might be available even though exhaustion operates to eliminate patent damages”). *Id.* at 637. *See also* *Motion Picture Patents Co. v. Universal Film Mfg. Co.*, 243 U.S. 502 (1917) (“The extent to which the use of the patented machine may validly be

ease of movement of both software and hard goods around the world, this ambiguity impacts on a global level.¹¹

Recently, and perhaps counterintuitively, some antitrust scholars and other commentators have argued that the economics of post-sale restraints provide a robust theoretical foundation for undermining exhaustion rules or at least narrowing their scope, suggesting that just as antitrust law has relaxed its previous animosity towards post-sale restraints, IP law should embrace them as well.¹²

In this Article, I argue that the economics of post-sale restraints do not support the case against exhaustion rules. While these insights can help fill some of the doctrine's missing theoretical foundations and assist in drawing its proper contours, they do not support the case against the first sale doctrine. Instead, these insights ultimately support the doctrine's validity and continued vitality. This Article shows that the economic literature teaches us that (a) the doctrine should not be relied on to automatically invalidate short-term *contractual* post-sale restraints, and (b) that the law might actually uphold such contracts when they are entered into between collaborating and non-integrated firms when they are necessary to encourage investment and control opportunism.¹³ But the Article also shows that the economic literature also teaches us that firms

restricted to specific supplies or otherwise by special contract between the owner of a patent and the purchaser or licensee is a question outside the patent law, and with it we are not here concerned.") *Id.* at 509.

11. *See, e.g.*, Trade-Related Aspects of Intellectual Property Rights ("TRIPS"), Marrakesh Agreement Establishing the World Trade Organization, Apr. 15, 1994, 1867 U.N.T.S. 154, Annex 1C, art. 6 (agreement to disagree on the issue of exhaustion); *Euro-Excellence Inc. v. Kraft Canada Inc.*, [2007] 3 S.C.R. 20 (Can.). Abraham Drassinower writes that the *Euro-Excellence* case "regales us not with one or two, but with nothing less than four different judgments The effect of these overlapping yet distinct and concurring judgments, which both partially agree and partially disagree with each other in multiple respects, is that, aside from the relatively easy statement that the defendant parallel importer won the case, it is difficult to identify with clarity or conviction what the law of parallel imports of copyrighted works is in Canada." Abraham Drassinower, *The Art of Selling Chocolate: Remarks on Copyright's Domain*, in FROM "RADICAL EXTREMISM" TO "BALANCED COPYRIGHT": CANADIAN COPYRIGHT AND THE DIGITAL AGENDA 121, 122 (Michael Geist ed., 2010).

12. *See, e.g.*, *The First Sale Doctrine*, INTELLECTUAL PROPERTY COLLOQUIUM (Mar. 2010), <http://www.ipcolloquium.com/Programs/12.html> (discussing the first sale doctrine and comparing copyright's approach to that of antitrust law).

13. *See infra* Part VI.

tend to consider the short term but discount the long term, and that in calculating the costs and benefits of their activities, they consider only their own private costs and benefits, but not those external to them.¹⁴ Such externalities and other market imperfections strongly imply that even when IP owners seek to impose post-sale restraints for economically justified reasons, the nature, scope, and duration of those restraints would be socially suboptimal. Thus, the first sale doctrine permits IP owners to implement and enforce some post-sale restraints when they are most likely to be beneficial—during the production or initial distribution stages of their products—but otherwise guarantees that the use and reuse of such products would not be burdened with excessive and unnecessary restraints. Taken seriously, these economic insights suggest that the doctrine should remain a sticky default *property* rule in most other circumstances.

This Article focuses on the first sale doctrine in copyright law, but most of the analysis is applicable to other IP rights.¹⁵ The Article is organized as follows: Part II describes different existing and possible formulations of the first sale doctrine to set the stage for the discussion that follows. Part III briefly traces the origin of the first sale doctrine in copyright law and reminds us of the role that unexhausted IP rights can play in facilitating anticompetitive practices such as industry-wide cartels; Part IV describes how modern antitrust law treats vertical restraints and how this view can inform the debate surrounding the first sale doctrine. The Article then discusses how careless application of the insights derived from modern antitrust law and economics has misinformed several debates. Parts V discusses the misapplication of these insights in the context of parallel trade, Part VI criticizes the insufficient attention the distinction between property rights and those arising out of contracts, and Part VII shows more generally how a hurried application of the Coase theorem misinforms many of these and related debates. Part VIII considers the continuing vitality of the

14. See *infra* Part VII.

15. Indeed, since the first sale doctrine is merely a manifestation of “a common-law doctrine with an impeccable historic pedigree,” *Kirtsaeng v. John Wiley & Sons, Inc.*, 133 S. Ct. 1351, 1363 (2013), the analysis can apply even more broadly to all restraints on alienation of property, whether or not IP rights are involved.

doctrine; and Part IX revisits several recent cases in light of the insights of this Article.

II. THE FIRST SALE DOCTRINE: DIFFERENT FORMULATIONS

Part of the difficulty in the debate about the first sale doctrine is the fact that its rationale has never been fully articulated, nor have its boundaries been made clear. Unfortunately, despite more than a hundred years of litigation, many questions linger on, without clear or satisfactory answers. Moreover, even when the doctrine applies, it may be unclear how strong it is, that is, whether the first sale doctrine is a mandatory rule or merely a default rule, and if it is a default rule, how sticky the default rule is. In order to begin answering these questions, it may be useful to recognize that there might be at least five possible formulations of the doctrine based on perceived strength.

A. No First Sale Doctrine

At one end of the spectrum lies the option of no first sale doctrine, meaning that it is up to the IP owner to decide whether the first-authorized sale or any subsequent sale would exhaust the IP right. I include this option on the spectrum not merely for the elegance of the model, but also because in some instances this is (or might be) the law. For example, if, as the Ninth Circuit held in *Vernor*,¹⁶ the doctrine does not apply when a software publisher chooses to characterize a transaction as a license rather than a sale, then every subsequent sale or disposition of the copy requires the copyright owner's permission. If, as the district court held in *ReDigi*,¹⁷ the first sale doctrine does not apply to digital copies resold apart from the particular medium in which they are embedded, then a large and growing swath of markets for digitally distributed works continue to be under the constant control of copyright owners. And had the dissent in *Kirtsaeng* had its way, the first sale doctrine would apply only to copies made in the United States, but not to copies made abroad.¹⁸ Had this been the case,

16. *Vernor v. Autodesk, Inc.*, 621 F.3d 1102 (9th Cir. 2010).

17. *Capitol Records, LLC v. ReDigi Inc.*, 934 F. Supp. 2d 640 (S.D.N.Y. 2013).

18. *Kirtsaeng*, 133 S. Ct. at 1376 (Ginsburg, J., dissenting); *Quality King Distribs., Inc.*

copyright owners could avoid the first sale doctrine by locating production offshore,¹⁹ meaning that the copyright owner could effectively dictate whether the doctrine applies or not.

B. Weak First Sale Doctrine

A weak formulation views the first sale doctrine as a simple default rule: the IP owner's exclusive right is exhausted after the first *unconditional* sale. However, it may not be exhausted if the transaction is conditional and conditions were imposed by license, contract, or, possibly, even mere notice. Breach of any such conditions (by a contracting party, a licensee, or a purchaser who had notice of the restriction) would trigger liability for infringement, subject to antitrust scrutiny or an antitrust-based IP misuse defense, placing the burden of showing an anticompetitive effect or misuse on the person who challenges the restraint.²⁰

C. Moderate First Sale Doctrine

Under a moderate first sale doctrine the first authorized sale still exhausts the IP right but the buyer may still be bound by contractual post-sale restraints. Under this reading, contractual post-sale restraints are valid (subject to general antitrust or misuse scrutiny), yet their breach may trigger liability only for breach of contract, rather than for IP infringement. The consequences are that only weaker remedies for breach contract but not the stronger IP remedies will be available, and that those remedies will be available against a narrower set of parties (those privy to the contract, but not others).²¹

v. L'anza Res. Int'l, Inc., 523 U.S. 135, 154 (1998) (Ginsburg, J., concurring); *see also* BMG Music v. Perez, 952 F.2d 318 (9th Cir. 1991).

19. *But see* Parfums Givenchy, Inc. v. Drug Emporium, Inc., 38 F.3d 477, 481 (9th Cir. 1994) (holding that the while the first sale doctrine does not apply to copies made outside the United States, it does apply upon the first authorized sale of those copies in the United States).

20. *See, e.g.*, Herbert Hovenkamp, *Post-Sale Restraints and Competitive Harm: The First Sale Doctrine in Perspective*, 66 N.Y.U. ANN. SURV. AM. L. 487, 541 (2010).

21. This was the position of the United States in its Amicus Brief in *Quanta Computer, Inc. v. LG Electronics*. *See* Brief of Amicus Curiae United States in Support of Petitioners, *Quanta Computer, Inc. v. LG Elecs.*, 553 U.S. 617 (2008) (No. 06-937), 2007 WL 3353102, at *29 ("Some of the same restrictions that the first sale doctrine renders ineffective in a patent-infringement suit could be validly imposed as a matter of state contract law. But

D. Strong First Sale Doctrine

A strong formulation of the rule treats the first sale doctrine as a sticky default rule de jure.²² This means that the first authorized sale exhausting the IP rights and attempting to work around exhaustion rules would be invalidated in the absence of a compelling case-specific explanation as to why the work around should be upheld. Procedurally, this formulation may translate into a rebuttable presumption of invalidity attached to attempts to work around the doctrine. The strength of such a presumption (or the ease of rebutting it) may evolve as, through adjudication, the case law may develop categories of cases of more or less permissible work-arounds.²³

E. Strongest First Sale Doctrine

Under the strongest formulation of the first sale doctrine, the sale exhausts the right, and attempts to work around the doctrine would be preempted, held invalid, may constitute IP misuse per se (i.e., without any inquiry into whether working around the doctrine can be justified),²⁴ or, at the extreme, trigger antitrust liability.

With these different possible formulations in mind, let me turn to discussing the origins of the doctrine and how the insights from modern antitrust law and the thinking about the economics of post-sale restraints help in choosing the right formulation.

even otherwise valid contract provisions would not provide a defense to a federal antitrust action, and mere unilateral notice to downstream purchasers will not generally give rise to enforceable contractual restrictions.”).

22. On sticky default rules, see generally Cass R. Sunstein, *Switching the Default Rule*, 77 N.Y.U. L. REV. 106 (2002). While default rules often tend to be sticky, for the reason discussed in Part VII below, the first sale doctrine may not be a sticky default unless its stickiness is bolstered legally.

23. This can be analogous to how the scope of fair use in copyright law develops, see Peter DiCola & Matthew Sag, *An Information-Gathering Approach to Copyright Policy*, 34 CARD. L. REV. 173 (2012), or to how antitrust jurisprudence moves between rule of reason and per se rules.

24. See *Omega S.A. v. Costco Wholesale Corp.*, No. CV04-05443TJH, 2011 WL 8492716 (C.D. Cal. Nov. 9, 2011) (holding that Omega’s attempt to prevent parallel importation of noncopyrightable watches by suing the importer for unlawfully importing the copyrighted company logo embedded in the back of the watch constituted copyright misuse).

III. THE ORIGINS

*Bobbs-Merrill v. Straus*²⁵ is often considered to be the grand entry of the first sale doctrine into American copyright law. A similar principle of exhaustion in patent law was reconfirmed shortly thereafter.²⁶ But like many legal rules, its genesis can be traced to more than a single source or event. As others have noted, the first sale doctrine emerged from an era in which IP rights were more modest in scope²⁷ and greater regard was given to their limited nature.²⁸ One could argue that it was formed in a more formalist era when legal thinking and reasoning tended to appreciate the distinction between legal categories and were averse to blurring the lines between them.

However, the main legal streams that gave rise to the first sale doctrine were the common law hostility to restraints on alienation, to servitudes in general, and servitudes in chattels in particular, and the emergence of antitrust law and its contemporaneous hostility towards both horizontal and vertical restraints.²⁹ Both of these streams have gone through substantive transformation over time. The common law's aversion to servitudes has waned, and antitrust law, influenced by modern antitrust scholarship, has not only relaxed its hostility toward vertical restraints, but in some cases has grown to

25. *Bobbs-Merrill Co. v. Straus*, 210 U.S. 339 (1908).

26. *See* *Bauer & Cie. v. O'Donnell*, 229 U.S. 1 (1913); *Motion Picture Patents Co. v. Universal Film Mfg.*, 243 U.S. 502 (1917); *see also* Katherine J. Strandburg, *Users as Innovators: Implications for Patent Doctrine*, 79 U. COLO. L. REV. 467, 495 (2008) ("The first sale doctrine holds that a patentee's rights are 'exhausted' when a patented product is sold, leaving the purchaser free to do with it as he or she wishes.").

27. *See, e.g.*, Neil Weinstock Netanel, *Copyright and a Democratic Civil Society*, 106 YALE L.J. 283, 294 (1996) (characterizing modern copyright law as "bloated"); *see also* Lionel Bentley, *R. v The Author: From Death Penalty to Community Service*, 32 COLUM. J.L. & ARTS 1, 3 (2008) (describing the consensus among academic about the expansion of copyright).

28. *Bobbs-Merrill*, 210 U.S. at 346 ("The copyright statutes ought to be reasonably construed, with a view to effecting the purposes intended by Congress. They ought not to be unduly extended by judicial construction to include privileges not intended to be conferred, nor so narrowly construed as to deprive those entitled to their benefit of the rights Congress intended to grant."); *see also* *Motion Picture Patents*, 243 U.S. at 510.

29. Hovenkamp, *supra* note 20, at 504 ("Over history, most of the Supreme Court's decisions on the first sale doctrine have attached its rationale to competition policy. . . . Lacking a rationale for explaining why vertical restrictions were anticompetitive in the traditional sense of leading to reduced output and higher prices, antitrust itself imported from the first sale doctrine the common law's concern with restraints on alienation.").

embrace them enthusiastically. Some may argue that just as the confluence of property law and antitrust gave rise to the first sale doctrine, a new confluence of these streams flowing in the reverse direction will lead to its demise.

A. *The Antitrust Legacy of Bobbs-Merrill*

Bobbs-Merrill v. Straus, decided in 1908, is often described in antitrust literature as a resale price maintenance (RPM) case,³⁰ namely, an attempt by a manufacturer (a publisher) to maintain the resale price of her product (a book). Bobbs-Merrill was the publisher of a novel titled *The Castaway* by Hallie Ermine Rives.³¹ A notice printed on the book prohibited anyone from reselling it for less than \$1 and threatened to treat any violation of this condition as copyright infringement. Macy's, then a discount department store, purchased copies of the book from a distributor and sold them for eighty-nine cents each.³² Bobbs-Merrill sued for copyright infringement. The Supreme Court dismissed the action, holding that the publisher's exclusive right to sell copies does not give it the power to "qualify the title of a future purchaser To add to the right of exclusive sale the authority to control all future retail sales, by a notice that such sales must be made at a fixed sum, would give a right not included in the terms of the statute, and, in our view, extend its operation, by construction, beyond its meaning, when interpreted with a view to ascertaining the legislative intent in its enactment."³³

Strictly speaking, *Bobbs-Merrill* was not an antitrust case. Bobbs-Merrill sued for copyright infringement, and the Court's decision focused on the scope of a copyright owner's exclusive right to vend copies.³⁴ However, antitrust scholarship often treats the decision as a reflection of the Court's contemporaneous hostility to all types of post-sale restraints, whether based in property (and intellectual

30. *Id.* at 497; *see also* KEITH N. HYLTON, *ANTITRUST LAW: ECONOMIC THEORY AND COMMON LAW EVOLUTION* 261 (2003).

31. HALLIE ERMINE RIVES, *THE CASTAWAY* (1904).

32. *Bobbs-Merrill*, 210 U.S. at 340–42.

33. *Id.* at 351.

34. *Id.*

property) rights or in contract.³⁵ Indeed, three years after *Bobbs-Merrill*, the Court expressed similar hostility, this time as a matter of antitrust law, when it held in *Dr. Miles Med. Co. v. John D. Park & Sons Co.*³⁶ that resale price maintenance was illegal per se under the antitrust laws.³⁷ According to Hovenkamp, this trend culminated in a duo of cases, *Straus v. Victor Talking Machine* and the *Motion Picture Patents* case, both issued on the same day in April 1917, which effectively created a merger between the first sale doctrine and antitrust policy.³⁸

Viewed from this perspective, the first sale doctrine and antitrust law's treatment of vertical restraints are nothing but two sides of the same coin. One reflects aversion to relying on IP rights to impose such restraints, the other manifests hostility to instituting them by contract. If so, then change in the law's attitude towards one side of the coin might appear necessarily to require change in the other. The shortcomings of this logic will be discussed in greater length in Parts VI and VII. At this point, however, it will be worthwhile to consider a less known aspect of *Bobbs-Merrill's* antitrust history, because the lessons from this unfamiliar story will also be relevant for appreciating the merits of the first sale doctrine and the perils of abandoning it.

The Court's decision in the *Bobbs-Merrill* case ends with a remark explaining that in light of the Court's conclusion to dismiss the copyright infringement claim, "it [is] unnecessary to discuss other questions noticed in the opinion in the circuit court of appeals, or to examine into the validity of the publisher's agreements, alleged to be in violation of the acts to restrain combinations creating a monopoly or directly tending to the restraint of trade."³⁹

This single allusion to antitrust opens the door to a much richer antitrust history discussed in the lower courts' decision. It turns out that the dispute in *Bobbs-Merrill* did not involve a single publisher

35. Hovenkamp, *supra* note 20, at 503–10.

36. *Dr. Miles Med. Co. v. John D. Park & Sons Co.*, 220 U.S. 373 (1911).

37. *Id.* at 408, *overruled by* *Leegin Creative Leather Prods., Inc. v. PSKS, Inc.*, 551 U.S. 877 (2007) (holding that resale price maintenance is subject to rule of reason analysis rather than per se illegality).

38. Hovenkamp, *supra* note 20, at 508.

39. *Bobbs-Merrill*, 210 U.S. at 351.

desiring to exercise control over the distribution system of its books. Rather, the case concerned an industry-wide cartel established in 1900 comprising the majority of book publishers and booksellers who conspired to “correct some of the evils connected with the cutting of prices on copyright books” by requiring all publishers to fix the retail prices of their respective books to prevent discounting and eliminate discounters by collectively refusing to deal with them.⁴⁰ RPM was the means to enforce the cartel, and copyright was an effective method to enforce the RPM.

In fact, as early as 1903, Macy’s attempted to enjoin the cartel by suing the American Publishers’ Association and the American Booksellers’ Association under the New York Anti-Monopoly Act.⁴¹ Macy’s bid was only partially successful. The New York Court of Appeals found in favor of Macy’s on the basis that the publishers’ and booksellers’ rules sought to prevent the discounting of all books, copyrighted and noncopyrighted alike.⁴² Relying on *Bement v. National Harrow*,⁴³ however, the court held that a combination to fix the prices of copyrighted books alone would be legal. Soon thereafter, the respective associations’ rules were amended to cover copyrighted books only,⁴⁴ and Bobbs-Merrill commenced its copyright infringement suit against Macy’s. In its defense, Macy’s argued that the notice included in the book, which was lawfully printed and sold, could not restrict the right of the book’s owner to resell it as the owner saw fit. Macy’s further defended on the grounds that the lawsuit was an attempt to enforce an unlawful combination.⁴⁵ The Circuit Court sided with Macy’s on both issues

40. *Bobbs-Merrill Co. v. Straus*, 139 F. 155, 157 (C.C.S.D.N.Y. 1905), *aff’d*, 147 F. 15 (2d Cir. 1906), *aff’d*, 210 U.S. 339 (1908).

41. *Straus v. Am. Publishers’ Ass’n*, 69 N.E. 1107 (N.Y. 1904).

42. *Id.* at 477–81.

43. *E. Bement & Sons v. Nat’l Harrow Co.*, 186 U.S. 70, 91 (1902) (“[T]he general rule is absolute freedom in the use or sale of rights under the patent laws of the United States. The very object of these laws is monopoly, and the rule is, with few exceptions, that any conditions which are not in their very nature illegal with regard to this kind of property, imposed by the patentee and agreed to by the licensee for the right to manufacture or use or sell the article, will be upheld by the courts. The fact that the conditions in the contracts keep up the monopoly or fix prices does not render them illegal.”).

44. *Bobbs-Merrill*, 139 F. at 172.

45. *Id.* at 156.

and refused to construe *Bement v. National Harrow* as broadly as the state court did.⁴⁶ On appeal, the Second Circuit affirmed, but addressed only the copyright question.⁴⁷ The Supreme Court, as noted above, mentioned the antitrust issue only in passing and based its decision solely on copyright grounds.

This antitrust history of *Bobbs-Merrill* illustrates an important lesson about the merits of the first sale doctrine and the perils of abandoning it: there is an important difference between obligations enforced as property and those that could be enforced only by contract; the relative strength of property entitlements over contractual ones may be an advantage when they promote socially beneficial goals but a serious problem when they can be relied on to enforce harmful practices. It is well known that while cartels can be highly beneficial to their members (and detrimental to consumers), they are unstable because each of their members can do even better by cheating. In order to be sustainable, cartel members need to be able to monitor cheating and enforce adherence to cartel rules, and it has been well documented that RPM can be an effective means to enforce cartels; if retail prices are fixed, cheating may be easily monitored and the colluding firms may demand that the deviant retailer be disciplined.⁴⁸

While cartels may employ contractual RPM to stabilize or enforce their rules, the utility of RPM is limited because it depends on the ability to enforce the contractual obligation, and the obligation will not be binding on third parties. *Bobbs-Merrill* is a case on point. Macy's was not a party to the price fixing agreements and therefore was not bound by them. Even though it had been blacklisted by the publisher/bookseller cartel for being a price-cutter, Macy's still managed to procure copies of the book and offer them for sale at a lower price. The copyright infringement claim, then, was designed to

46. *Id.* at 193.

47. *Bobbs-Merrill Co. v. Straus*, 147 F. 15 (2d Cir. 1906).

48. See Ittai Paldor, *The Vertical Restraints Paradox*, 58 U. TORONTO L.J. 317, 328 (2008); see also Edward Iacobucci, *The Case for Prohibiting Resale Price Maintenance*, 19 WORLD COMP. L. & ECON. REV. 71 (1995).

force Macy's to comply with the cartel's rules, by purporting to enforce those terms against a third party.⁴⁹

The Supreme Court's holding on copyright grounds destabilized the cartel because it meant that Macy's could, as a matter of law, continue to cut prices. Therefore, it could be argued that the kind of problem that the Court faced in *Bobbs-Merrill* hardly justifies the first sale doctrine because if the problem were an underlying cartel, antitrust law would have no problem condemning it. Nevertheless, the case still teaches an important lesson. While cartels are illegal, breaking them requires action and their outright condemnation demands evidence of their existence. Moreover, in oligopolistic markets, firms may tacitly collude and lawfully coordinate their behavior to achieve a cartel-like outcome without any illegal agreement.⁵⁰ The fact that many IP-based markets are made up of a small number of large competitors,⁵¹ and that virtually any product can be sold accompanied by a trademarked and often copyrighted logo,⁵² should give rise to a concern that unexhausted IP rights could be used to facilitate tacit collusion in such markets. Thus, the first sale doctrine is the failsafe against these invisible cartels.

The lesson from *Bobbs-Merrill* can be summarized in the following proposition:

49. *Bobbs-Merrill*, 139 F. at 178 (“It . . . is found as a fact that such notice was put in such books, and that its enforcement as an alleged license agreement is . . . an attempt by complainant, as a member of said American Publishers’ Association, to enforce as against this defendant the rules of such associations and combination fixing prices, in an effort to maintain them. It is part of a scheme, and the right of the complainant to maintain this action depends on the validity of that scheme or combination.”).

50. See *Brooke Grp. Ltd. v. Brown & Williamson Tobacco Corp.*, 509 U.S. 209, 227 (1993) (“Tacit collusion, sometimes called oligopolistic price coordination or conscious parallelism, describes the process, not in itself unlawful, by which firms in a concentrated market might in effect share monopoly power, setting their prices at a profit-maximizing, supracompetitive level by recognizing their shared economic interests and their interdependence with respect to price and output decisions.”).

51. See Ariel Katz, *Making Sense of Nonsense: Intellectual Property, Antitrust, and Market Power*, 49 ARIZ. L. REV. 837, 866 (2007) [hereinafter Katz, *Making Sense*].

52. See, e.g., *Costco Wholesale Corp. v. Omega, S.A.*, 131 S. Ct. 565 (2010) (the alleged copyright infringement was the importation of watches bearing an engraved copyrighted logo); *Euro-Excellence Inc. v. Kraft Canada Inc.*, [2007] 3 S.C.R. 20 (Can.) (the alleged copyright infringement was the importation of chocolate bars bearing copyrighted logos on their wrappers).

Proposition 1: Beware of unexhausted IP rights

The antitrust history of *Bobbs-Merrill* teaches an important lesson: post-sale restraints that are enforceable with unexhausted IP rights can facilitate cartels and other types of anticompetitive behavior better than contractual restraints. A broader lesson, which will be further developed as the Article proceeds, is that while post-sale restraints may be efficient in certain circumstances, careful attention should be given to the question of what legal instruments should be used to enforce them. Put simply, with great power comes great responsibility; we should be wary of handing the great power of property entitlements to irresponsible hands.

IV. MODERN ANTITRUST APPROACH TOWARDS POST-SALE
RESTRAINTS

Post-sale restraints are a subset of a broader type of restraints, known as vertical restraints. Antitrust law's attitude towards agreements imposing such restraints (e.g., between manufacturers and distributors) has seen remarkable changes throughout its history.

Since it was enacted in 1890, Section 1 of the Sherman Act has declared illegal and punishable “[e]very contract, combination in the form of trust or otherwise, or conspiracy, in restraint of trade or commerce among the several States, or with foreign nations.”⁵³ It soon became common ground, however, that although “*every* contract, etc.” may be subject to scrutiny, not *all* contracts are equal. One fundamental distinction is between horizontal agreements, or those entered between actual or potential competitors (i.e., firms at the same level of the distribution chain), and vertical agreements, or those entered between firms at different levels of the distribution chain (e.g., manufacturer-wholesaler; wholesaler-retailer; retailer-customer). Another distinction is between restraints that are deemed *per se* illegal and those subject to rule of reason analysis, meaning that they would not be deemed illegal unless there is proof of anticompetitive effect. “Resort to *per se* rules is confined to restraints . . . ‘that would always or almost always tend to restrict

53. 15 U.S.C. § 1 (2012).

competition and decrease output.”⁵⁴ Further, to justify a per se prohibition a restraint must have “manifestly anticompetitive” effects and “lack . . . any redeeming virtue.”⁵⁵

Early antitrust law was as averse to vertical restraints as it was hostile to horizontal restraints, and treated many agreements imposing vertical restraints, such as exclusive dealing, tying, and resale price maintenance, as per se illegal.⁵⁶ Over time, however, antitrust scholarship began to recommend that horizontal and vertical restraints be treated differently because vertical restraints are more likely used to enhance efficiency, prevent opportunism, and otherwise advance procompetitive outcomes than they are implemented for anticompetitive ends.⁵⁷ Jurisprudence followed scholarship, and since the late 1970s,⁵⁸ the Supreme Court has gradually abolished virtually all per se rules applying to vertical restraints.⁵⁹ The last bastion fell in 2007 in *Leegin*, which overruled an almost century-old per se prohibition on resale price maintenance.⁶⁰ Abolishing per se prohibitions reflects the prevailing wisdom that not only are such restraints not necessarily harmful, but frequently they are beneficial.

The opening up toward vertical restraints has not crossed over to horizontal restraints because horizontal restraints are perceived to be more harmful than vertical restraints, and with good reason. Horizontal agreements have the ability to eliminate competition between rival firms, and this, in turn, results in output limits, raised prices, and monopoly profits.⁶¹ The resulting harm is not limited to the higher prices that consumers have to pay and the resulting

54. *Leegin Creative Leather Prods., Inc. v. PSKS, Inc.*, 551 U.S. 877, 886 (2007) (citation omitted).

55. *Id.*

56. *Id.* at 888.

57. *Id.* at 889.

58. *Continental T. V., Inc. v. GTE Sylvania Inc.*, 433 U.S. 36 (1977).

59. *Leegin*, 551 U.S. at 900–04 (2007).

60. A nominally per se rule against tying still exists, but it has been watered down so significantly that it is questionable whether it is different from a rule of reason. See Katz, *Making Sense*, *supra* note 51, at 896.

61. Frank H. Easterbrook, *Vertical Agreements and the Rule of Reason*, 53 ANTITRUST L.J. 135, 140 (1984) (“The antitrust laws exist to stop trusts or cartels from reducing the output of goods and services. The reduction brings about monopoly prices, the conspirators’ goal.”).

transfer of surplus from consumers to producers; “[i]t also causes a misallocation of resources, as people shift to things that seem ‘cheaper’ but really cost society more to produce in order to get the same level of satisfaction.”⁶²

In contrast, vertically situated parties do not share a common interest in reduced competition. Quite the contrary, a manufacturer benefits both when competition among its suppliers is intense and when market power among its distributors is limited.⁶³ Similarly, distributors would rather be free to deal with competing manufacturers upstream and competing retailers downstream than with monopolistic firms. In other words, the divergence of interests inherent in vertical relationships serves as a check on anticompetitive practices.⁶⁴

Thus, while it is possible that vertical restraints can be used anticompetitively, a growing body of literature has identified myriad reasons supporting the proposition that they often serve procompetitive goals.⁶⁵ For example, vertical restraints may encourage dealers to invest in developing a local market (by advertising or other means) or to supply pre- or post-sale services (such as training or repairs). I will discuss some of these procompetitive explanations in Part V below.

Post-sale restraints⁶⁶ may restrict what a buyer can do with the goods she purchased—where she can resell them, to whom, at what

62. *Id.*

63. HYLTON, *supra* note 30, at 253 (“Lower prices at the retail level enhance the manufacturer’s revenue, other things being equal. That is, once the manufacturer sells the good to the retailer, it would prefer to see the retailer charge the lowest possible price, since that maximizes the quantity sold. Indeed, once the manufacturer has sold his output to retailers, he would be happy to see them drive the resale price to zero.”).

64. *Leegin*, 551 U.S. at 896 (“[I]n general, the interests of manufacturers and consumers are aligned with respect to retailer profit margins. The difference between the price a manufacturer charges retailers and the price retailers charge consumers represents part of the manufacturer’s cost of distribution, which, like any other cost, the manufacturer usually desires to minimize.”).

65. *See, e.g.*, MICHAEL J. TREBILCOCK ET AL., THE LAW AND ECONOMICS OF CANADIAN COMPETITION POLICY 373–99 (2003) (surveying the various pro- and anti-competitive theories of vertical restraints).

66. Like Hovenkamp, I use the term “post-sale restraint” to refer “generically to any restriction imposed by a seller on how a purchased good can be used or resold after the initial sale” and in this context the term “sale” includes leases, licenses, or other transfers of interest short of a technical sale.” *See* Hovenkamp, *supra* note 20, at 487.

prices, and whether she will have to provide pre- or post-sale services, repairs, warranties, etc. As modern antitrust law has opened up to vertical restraints and their positive impact on competition, it has grown to tolerate agreements that govern and monitor them. In the case of goods in which no IP rights subsist, contracts and the mere threat of termination serve as the main tools for enforcing such restraints.⁶⁷ In contrast, if IP rights can be relied on to enforce the restraints, the additional set of remedies could make those restraints more effective.⁶⁸

Notably, remedies for breach of contract are usually weaker than remedies available for infringement of an intellectual property right. Remedies for breach of contract are often limited to expectation damages and generally would be filed in state courts.⁶⁹ For example, even though the IP infringement case of *eBay v. MercExchange*⁷⁰ made the grant of injunctive relief less certain than before, injunctions are still more likely to be granted in IP infringement cases than for breach of contract. The threat of injunction (especially when the user faces high switching costs),⁷¹ coupled with statutory damages that might be available in copyright,⁷² or treble damages in the case of willful patent infringement,⁷³ and attorneys' fees in both, may better ensure compliance with the imposed restrictions than contracts. Moreover, contractual restrictions are enforceable only against the purchaser who is a party to the contract, but not against third parties such as the purchaser's customers, users, or service personnel.⁷⁴ Therefore, if IP rights subsist in the goods and can be

67. Contract law is the main, but not the exclusive tool. On some occasions, plaintiffs were able to enforce the restraints against third parties relying on some common law commercial torts or even concepts of equitable servitudes created by notice. *See, e.g.*, *Clairol, Inc. v. Sarann Co.*, 37 Pa. D. & C.2d 433 (1965).

68. Hovenkamp, *supra* note 20, at 492.

69. *Id.* at 539.

70. *eBay, Inc. v. MercExchange, L.L.C.*, 547 U.S. 388 (2006).

71. Ariel Katz, *A Network Effects Perspective on Software Piracy*, 55 U. TORONTO L.J. 155, 207 (2005) [hereinafter Katz, *Network Effects*].

72. 17 U.S.C. § 504(c).

73. Christopher A. Harkins, *A Budding Theory of Willful Patent Infringement: Orange Books, Colored Pills, and Greener Verdicts*, 2007 DUKE L. & TECH. REV., no. 6, at 1, 6.

74. Hovenkamp, *supra* note 20, at 541.

relied on to enforce post-sale restraints, producers have a more powerful and arguably more effective tool to enforce such restraints.

The insights from modern antitrust and the economics of post-sale restraints, namely, that post-sale restraints may be economically justified, support rejection of the strongest version of the first sale doctrine; yet, they do not easily help choosing which of the remaining versions is preferable. Still, many antitrust-minded discussions about the first sale doctrine tend to favor one of the weaker versions of the doctrine, ranging from preference to no exhaustion in the case of parallel imports, to weak or moderate versions in other contexts. The following parts of this Article will identify three central flaws in the way the insights of modern antitrust are relied on to justify the no/weak/moderate versions of first sale doctrine, as well as demonstrate why the preferable version is a strong one, namely treating the doctrine as a sticky default rule.

The first flaw arises in the context of debates on the issue of parallel trade, where the insights from modern antitrust are harnessed to support a rule of national rather than international exhaustion. As I explain below, this flaw is mainly a logical one, in the sense that, if valid, the logical conclusion from the argument is no exhaustion at all. The argument, however, fails to justify the purported distinction between national and international exhaustion. The second flaw occurs when the virtues of post-sale restraints are invoked to justify a weak version of the first sale doctrine while failing to notice the subtle, yet critical, difference between remedies for breach of contract and remedies for IP infringement, or embracing the notion of enhanced IP remedies without recognizing their associated costs. The failure here is to fully account for the critical institutional differences between property and contract. The third flaw, which essentially underlies the second, occurs when the first sale doctrine is treated merely as a default rule, assuming that working around it must be presumptively efficient. This approach is flawed because it adopts a Coasian logic in a notoriously non-Coasian setting, namely that of intellectual property.

V. THE FLAW IN THE PARALLEL TRADE DEBATE

The term “parallel trade” (or “gray market”) refers to situations in which goods sold abroad at a lower price than that charged locally are imported (or reimported) by an unauthorized dealer to be sold

domestically in competition with the same goods sold through the local authorized distribution system designated by the manufacturer.⁷⁵ When such goods embed some protected IP, IP owners sometimes invoke IP law to ban the unauthorized importation. Whether or not IP law should be used for this purpose is subject to a heated debate and the actual rules are often inconsistent among nations, as well as between different IP laws within a nation (i.e., different rules for patents, copyrights, and trademarks) and even within a specific national IP law.⁷⁶ One question that plagues this debate is whether the IP right is exhausted only upon the first *domestic* sale authorized by the IP owner (national exhaustion) or whether the IP right is exhausted upon the first sale authorized by the IP owner regardless of the country in which it occurs (international exhaustion). A third, hybrid rule mandates that the IP right is exhausted upon any first authorized sale within a regional free-trade area, but IP laws might still be used to block unauthorized imports from all other countries (regional exhaustion).⁷⁷ Regional exhaustion is currently the law in the European Economic Area (consisting of the European Union,

75. See *Kirtsaeng v. John Wiley & Sons, Inc.*, 133 S. Ct. 1351, 1379 n.9 (2013) (“The term ‘gray market good’ refers to a good that is ‘imported outside the distribution channels that have been contractually negotiated by the intellectual property owner.’”) (citing MIRANDA FORSYTH & WARWICK A. ROTHNIE, *PARALLEL IMPORTS, in THE INTERFACE BETWEEN INTELLECTUAL PROPERTY RIGHTS AND COMPETITION POLICY* 429 (S. Anderman ed., 2007)).

76. See, e.g., *Euro-Excellence Inc. v. Kraft Canada Inc.*, [2007] 3 S.C.R. 20 (Can.). It is interesting to compare the Justices’ reasoning in this case. Justice Rothstein held that parallel imports of a copyrighted work infringes the copyright only when the local copyright holder is an assignee rather than an exclusive licensee. *Id.* at 21–22. Justice Abella, on the other hand, rejected the distinction between an assignee and an exclusive licensee and held that the unauthorized importation infringes a copyright in both cases. *Id.* at 24–25 (Abella, J., dissenting). Justice Bastarache agreed with Justice Abella that the distinction between an assignee and exclusive licensee is immaterial, but he argued that the ban on importation would not exist when the work is a logo printed on a wrapper of a consumer good. *Id.* at 23–24 (Bastarache, J., concurring). For a thorough discussion of these different opinions, see Drassinower, *supra* note 11.

77. See *International Exhaustion and Parallel Importation*, WORLD INTELLECTUAL PROP. ASS’N, http://www.wipo.int/sme/en/ip_business/export/international_exhaustion.htm (last visited Jan. 13, 2014).

Iceland, Lichtenstein, and Norway).⁷⁸ Indeed, the difference between national exhaustion and international exhaustion was the dividing line between the majority and the dissenting opinions in *Kirtsaeng*.⁷⁹

Supporters of international exhaustion highlight the benefits of increased competition and lower prices that parallel imports can bring about and view IP-based prohibition on parallel trade as unjustified barriers to international trade.⁸⁰ Supporters of national exhaustion emphasize the benefits arising from tight control by producers over their distribution systems and argue that parallel imports undermine the integrity of such systems, which results in less efficient distribution systems to the detriment of producers and consumers alike.⁸¹

Arguments supporting a ban on parallel imports will be familiar to anyone acquainted with modern antitrust thinking about vertical restraints. The arguments typically warn that one or more of the advantages arising from vertical restraints would be lost without IP laws banning parallel imports. The discussion below considers these arguments and shows that they fail to support an IP-backed ban on parallel imports (i.e., national exhaustion).

A. Price Discrimination

Proponents of national exhaustion often cite the benefits of international price discrimination as a justification for banning parallel trade.⁸² Price discrimination occurs in this context when

78. Case C-355/96, *Silhouette Int'l Schmied GmbH & Co. KG v. Hartlauer Handelsgesellschaft mbH*, 1998 E.C.R. I-4799, 2 C.M.L.R. 953 (1998).

79. *Kirtsaeng*, 133 S.Ct. at 1384 (“[I]n my view, [section 602(a)(1)] ties the United States to a national-exhaustion framework. The Court’s decision, in contrast, places the United States solidly in the international-exhaustion camp.”) (Ginsburg, J., dissenting).

80. See, e.g., Frederick M. Abbott, *First Report (Final) to the Committee on International Trade Law of the International Law Association on the Subject of Parallel Importation*, 1998 J. INT’L ECON. L. 607, 622 (1998).

81. *Infra*, Parts V.A and V.B.

82. See *John Wiley & Sons, Inc. v. Kirtsaeng*, 654 F.3d 210, 221 (2d Cir. 2011), *rev’d*, 133 S. Ct. 1351 (2013) (reasoning that the first sale doctrine does not apply to copies made outside the United States because Congress “obviously intended to allow copyright holders some flexibility to divide or treat differently the international and domestic markets for the particular copyrighted work”).

demand for a product varies across countries and producers set different prices in different countries. In other words, producers may charge a higher price where demand (or ability to pay) is higher and a lower price where demand (or ability to pay) is lower. Parallel trade is a form of arbitrage that jeopardizes the producers' ability to price discriminate. Proponents of national exhaustion argue that parallel trade should be prohibited because price discrimination promotes consumer welfare.⁸³ The main benefits of price discrimination are suggested to be threefold: (1) it increases output,⁸⁴ thereby ameliorating some of the inefficiencies otherwise associated with the exclusivity inherent in IP rights⁸⁵ and benefiting consumers with lower ability to pay who would not be able to obtain the product if the seller could set only a uniform price;⁸⁶ (2) increased output may contribute to achieving economies of scale and learning resulting in lower per-unit costs;⁸⁷ and (3) even when output under price discrimination is lower compared to uniform pricing,⁸⁸ the additional

83. *Kirtsaeng*, 133 S. Ct. at 1390 n.27 (“It should not be overlooked that the ability to prevent importation of foreign made copies encourages copyright owners . . . to offer copies of their works at reduced prices to consumers in less developed countries who might otherwise be unable to afford them. The Court’s holding, however, prevents copyright owners from barring the importation of such low-priced copies into the United States, where they will compete with the higher priced editions copyright owners make available for sale in this country. To protect their profit margins in the U.S. market, copyright owners may raise prices in less developed countries or may withdraw from such markets altogether Such an outcome would disserve consumers—and especially students—in developing nations and would hardly advance the ‘American foreign policy goals’ of supporting education and economic development in such countries”) (Ginsburg, J., dissenting).

84. See, e.g., David A. Malueg & Marius Schwartz, *Parallel Imports, Demand Dispersion, and International Price Discrimination*, 37 J. INT’L ECON. 167 (1994).

85. *Id.* at 175–81, 189–90. *Contra* Frederick M. Abbott, *First Report (Final) to the Committee on International Trade Law of the International Law Association on the Subject of Parallel Importation*, 1 J. INT’L ECON. L. 607 at 619–21.

86. Malueg & Schwartz, *supra* note 84, at 190.

87. See Jerry A. Hausman and Jeffrey K. MacKie-Mason, *Price Discrimination and Patent Policy*, 19 RAND J. ECON. 253, 254 (1988).

88. As Guy Rub explains, “If the result of eliminating market segmentation is a cheap uniform price, then buyers in developed countries will be much better off, while buyers in developing countries will be slightly worse off (as it is almost certain that the uniform price will be at least somewhat more expensive than the price in developing countries under a well-functioning market segmentation scheme). Total buyers’ surplus will likely increase.” Guy A. Rub, *The Economics of Kirtsaeng v. John Wiley & Sons, Inc.: The Efficiency of a Balanced Approach to the First Sale Doctrine*, 81 FORDHAM L. REV. RES GESTAE 41, 46 (2013).

monopoly profit will encourage investment in R&D and innovation and contribute to dynamic efficiency.⁸⁹ In sum, the argument for price discrimination is that it promotes allocative efficiency (i.e., leads to higher output and/or lower costs) and dynamic efficiency (encourages innovation).

But in addition to highlighting the purported efficiency of price discrimination, proponents of national exhaustion sometimes tout it as socially just because it might help consumers with low willingness or ability to purchase goods, from which they would otherwise be excluded.⁹⁰ Thus, the first sale doctrine, which puts sand in the wheels of price discrimination, leads not only to deadweight loss, but also to distributional injustices.

These arguments sound compelling enough when discussing parallel importation and textbooks where, as it has been argued, international exhaustion might force publishers to “raise prices in less developed countries or may withdraw from such markets altogether.”⁹¹ How would a conscientious policy-maker adopt “[s]uch an outcome [that] would disserve consumers—and especially students—in developing nations and would hardly advance the ‘American foreign policy goals’ of supporting education and economic development in such countries”?⁹² But if textbooks are not convincing enough, then no other area demonstrates these points more dramatically than patented pharmaceuticals, where the affordability of drugs can be a matter of life or death. The high cost associated with developing new drugs is the main justification for allowing their patentability and for the high prices that are often charged for them, despite the inevitable limitations on accessibility.⁹³ As long as the development of new drugs is entrusted in the hands of profit-seeking private firms depending on patent protection, millions of people in developing countries inevitably will be priced out of the market. Arguably, price discrimination could ameliorate this problem

89. See, e.g., RICHARD POSNER, *ANTITRUST LAW* 203 (2d ed. 2001).

90. See, e.g., *Kirtsaeng v. John Wiley & Sons, Inc.*, 133 S. Ct. 1351, 1390 n.27 (2013) (Ginsburg, J., dissenting).

91. *Id.*

92. *Id.*

93. See, e.g., Alan O. Sykes, *TRIPS, Pharmaceuticals, Developing Countries, and the Doha “Solution”*, 3 CHI. J. INT’L L. 47, 57 (2002).

of accessibility and would immediately benefit both the manufacturer and the consumer in developing countries: drug companies could set high prices in developed countries and lower prices in less-developed ones. They could spread the cost of developing the drugs across all markets and make them available to a wider range of consumers.⁹⁴ It has been argued that drug companies would readily engage in price discrimination if they could prevent the cheap drugs from being re-imported into developed countries, but if they could not, they would rather not sell in the less-developed countries at all.⁹⁵ Banning parallel trade, therefore, not only seems to promote economic efficiency but also social justice.

However, while trade in pharmaceuticals vividly highlights the issue, it is somewhat of a red herring because some unique features of pharmaceuticals make them easy to be used as a stylized model, which is not easily generalizable to other goods.

First, pharmaceuticals are unique because the idea that people might die or suffer because they cannot afford the high cost of an available drug is morally offensive. If patents make an iPhone expensive, the fact that some consumers cannot afford it may be considered an unavoidable but acceptable outcome of our choice to encourage innovation by granting patents. However, in the case of pharmaceuticals, the deadweight loss that may result from a patent-related monopoly pricing becomes a life or death issue, which deserves closer attention.

Second, the case of patented drugs comes very close to being a stylized model favored by economists in the sense that it is relatively easy to assume that all else is equal and that differences in demand across countries will reflect only differences in ability to pay and not something else. We can assume that on average individuals' preference for their health and life is identical across countries⁹⁶ and that differences in demand reflect only differences in affordability (that is, the fact that a middle-class American is willing to pay for a

94. See, e.g., Patricia M. Danzon, *Price Discrimination for Pharmaceuticals: Welfare Effects in the US and the EU*, 4 INT'L J. ECON. BUS. 301 (1997).

95. See, e.g., Sykes, *supra* note 93, at 64.

96. Individuals may have different preferences (some are more health conscious than others; some have more desire to live than others), but we can assume that these differences are not strongly related to nationality, culture, or socio-economic conditions.

life-saving drug whereas a middle-class South-African may not pay for that same drug reflects only their different abilities to pay and not a difference in how each of them values the worth of living). We can also assume that the market structure is similar across nations in the sense that, all things being equal, a patented drug will have the same type and number of therapeutic substitutes in all countries.⁹⁷ If so, differences in demand across countries will not reflect differences in the cross-elasticity of demand between the drug and its substitutes within each country. In other words, variation in prices across countries will indeed reflect differences in ability to pay rather than differences in how competitive one market is compared to another.⁹⁸ The cost of developing new drugs and the truly global nature of the market makes compelling the proposition that price discrimination will enable patentees to better spread the cost globally and maximize the recovery of their investment.

Therefore, pharmaceuticals seem to present an example that may be close to a stylized model under which enacting laws banning parallel imports would improve welfare. Whether this is actually the case is another question. Let me raise some skepticism. First, trade in pharmaceuticals is strictly regulated. In the United States, for example, even if an astute arbitrageur identified an opportunity to import cheap drugs from a low-price country, pharmaceutical regulation laws would not allow it, irrespective of the exhaustion regime under patent law.⁹⁹ One could then assume that differential

97. It is possible, of course, that genetic differences could influence the amount of available substitutes within and across populations and that the future advances in pharmacogenomics would show such differences.

98. That is, if the patented drug is the only therapeutic remedy, its holder is a monopolist in all countries. If there are substitutes, they are available in all countries. This assumption might be too simplistic. Countries with small markets (either because they are small or because they have only a small number of residents who can afford paying) can be less competitive than larger markets because it may be more difficult to achieve economies of scale in distribution, marketing, regulatory compliance, etc. Prices in such cases may be higher.

99. Rebecca S. Eisenberg, *The Role of the FDA in Innovation Policy*, 13 MICH. TELECOMM. & TECH. L. REV. 345, 362–64 (2007). The EU is an exception. As part of the objective of creating a common market, the EU adopted a regional exhaustion regime. As a result, IP rights cannot be used to prevent the trade of pharmaceuticals between the member states. See Nicolas Petit, *Parallel Trade: Econ-oclast Thoughts on a Dogma of EU Competition Law*, in TRADE AND COMPETITION LAW IN THE EU AND BEYOND 332 (Inge Govaere et. al. eds., 2011).

pricing would be prevalent, and drugs would be universally affordable. This is not the case.¹⁰⁰ The fact that drugs are often still unaffordable in many developing countries despite the illegality in parallel trade suggests that there are other reasons for the lack of widespread price discrimination.¹⁰¹ Moreover, the theoretical argument would predict that developing countries—the supposed beneficiaries from bans on parallel imports¹⁰²—would support national exhaustion, but many of them strongly support parallel importation of drugs and perceive it as an important policy lever for increasing the affordability of medicines in their countries.¹⁰³ This suggests that even in a case that seems close to the stylized model, the reality may be much more complicated.¹⁰⁴

100. F. M. Scherer & Jayashree Watal, *Post-TRIPS Options for Access to Patented Medicines in Developing Nations*, 5 J. INTL. ECON. L. 913 (2002).

101. One reason is that many developing countries have a small population of highly affluent citizens and a majority of extremely poor citizens. In such cases, the local profit-maximizing price in the developing country is at a developed-country level, which eliminates the need for lower pricing. *Id.* at 930. Another reason might be fear of backlash in developed countries if legal versions were available cheaply abroad, resulting in the introduction of price controls (in the United States) or tighter price controls (in many other developed countries). *Id.*

102. *See, e.g.*, Sykes, *supra* note 93, at 64.

103. *See, e.g.*, CARLOS M. CORREA, WORLD HEALTH ORG., IMPLICATIONS OF THE DOHA DECLARATION ON THE TRIPS AGREEMENT AND PUBLIC HEALTH 17 (2002), available at <http://apps.who.int/medicinedocs/pdf/s2301e/s2301e.pdf>.

104. Textbooks may also seem like a textbook example of a case where national exhaustion would benefit developing countries. As Justice Ginsburg intimated, *Kirtsaeng v. John Wiley & Sons, Inc.*, 133 S. Ct. 1351, 1390 n.27 (2013) (Ginsburg, J., dissenting), national exhaustion would theoretically benefit developing countries by encouraging copyright owners to supply them with low-priced textbooks. While the facts in *Kirtsaeng* are consistent with this prediction, the reality is considerably more complicated. One might assume that since until the Court's holding in *Kirtsaeng*, copyright owners could prevent parallel importation of books, at least if they were printed abroad, scholars and students in developing countries would have enjoyed easy access to books at considerably lower prices. This does not seem to be the case. For example, Basheer et al. found in a sample of educational books in India (mainly in law and social sciences), international publishing houses tend to supply old and outdated editions, whereas the latest versions are available only through imports via websites (or through mainstream distributors), which cost as much as or more than their western counterparts. Shannad Basheer et al., *Exhausting Copyrights and Promoting Access to Education: An Empirical Take*, 17 J. INTELL. PROP. RTS. 335 (2012). While these findings may indicate the publishers refrain from supplying developing countries with the latest editions at low prices for fear of exportation, it is not clear how important the exhaustion regime is in developed countries. That is, arbitrage might have been a problem for publishers even prior to *Kirtsaeng*, but they have found ways to manage it, and would probably continue doing so even after

Unlike this stylized model, in most other cases differences in local demand may be the outcome of myriad reasons, with the result that it is impossible to generalize and determine a priori whether price discrimination should be encouraged or not.¹⁰⁵ For example, price differences may be a function of differences in the cross-elasticity of demand between the product and other products within each country. That is, consumers in countries *A* and *B* with the same disposable income might be willing to pay different prices for the same product because they have different preferences for the product given available alternatives. To illustrate further, consumers in Ontario might be willing to pay a higher price for ice-skates than consumers in Florida because they have fewer options for outdoor sports during the winter months (in economic terms, the cross-elasticity of demand between ice-skates and other sports equipment is high in Florida and low in Ontario). If trade is free and shipping costs are not too high, an ice-skates monopolist may not be able to price-discriminate by segregating the Ontario and Florida markets and may choose either to sell at a higher price in Ontario (and price consumers in Florida out of the market) or to set prices in both locations at the Florida level. We cannot determine a priori which outcome is more likely. Nor can we say a priori that maintaining the incentive to invest in developing better ice-skates requires an ability to segregate the Ontario and Florida markets.

Consider another possibility. It may turn out that because the Ontario market is large it attracts many ice-skate manufacturers and becomes very competitive. In contrast, the Florida market consists of a very small number of ice-skating devotees who are willing to pay high prices for skates, but is too small to sustain more than one brand. Under these circumstances, if the Florida market could be segregated from that of Ontario, consumers in Florida will be charged monopoly prices, whereas free trade will allow them to take advantage of the competition in Ontario by having skates shipped to them. Since we cannot tell with any degree of confidence that laws

Kirtsang. See also Rub, *supra* note 88, at 44–45 (noting that the outcome in *Kirtsang* is unlikely to eliminate price discrimination, only to prompt publishers to adopt other forms of price discrimination).

105. Generally, the welfare effect of price discrimination is ambiguous: it may or may not improve welfare. See generally TREBILCOCK ET AL., *supra* note 65, at 346.

prohibiting the importation of skates from Florida to Ontario or vice versa are preferable to free trade, there is no rational basis for enacting laws banning parallel trade.

As Nicolas Petit recently observed, price discrimination can also reflect different perceptions about a product's quality among consumers.¹⁰⁶ Petit provides the example of a bottle of Polish-branded vodka that "might be perceived as a special product in Western Europe—hence western EU consumers are ready to pay a high price for this product—and by contrast, be perceived as a relatively standard product in Poland—hence Polish consumers are only willing to pay a low price for this product. In a setting of this kind, the vodka producer will charge different prices in Poland and in western European countries."¹⁰⁷ Petit, who supports legal bans on parallel trade and exalts the benefits of price discrimination, maintains that there is "nothing intrinsically bad" about charging different prices in this situation.¹⁰⁸ Certainly the producer should not be morally or legally condemned for the fact that some consumers are willing to pay a premium for his brand, but there is nothing particularly laudable about such actions and it hardly follows that adopting laws bolstering this outcome is justified or desirable. If the better-informed Polish consumers are not willing to pay a premium for the particular brand, then we can assume that the price they are willing to pay is the more efficient one. Therefore, laws preventing the less-informed consumers from being offered the same price seem rather odd. Indeed, the ability of the marginal and better-informed consumer effectively to determine the market price that all consumers pay is one of the virtues of competitive markets.¹⁰⁹

106. Petit, *supra* note 99, at 336.

107. *Id.*

108. *Id.* (citing NAT'L ECON. RESEARCH ASSOCS. ET AL., THE ECONOMIC CONSEQUENCES OF THE CHOICE OF A REGIME OF EXHAUSTION IN THE AREA OF TRADEMARKS 104 (1998), available at http://ec.europa.eu/internal_market/indprop/docs/tm/report_en.pdf).

109. Douglas G. Baird, *The Boilerplate Puzzle*, 104 MICH. L. REV. 933, 936 (2005). And the ability of the less-informed consumer to piggyback on the knowledge of the better-informed depends on the inability of the seller to discriminate between them. See MARGARET JANE RADIN, *BOILERPLATE: THE FINE PRINT, VANISHING RIGHTS, AND THE RULE OF LAW* 104 (2012).

In sum, it is always possible to articulate a stylized model under which enacting laws banning parallel imports would improve welfare, and even if pharmaceuticals indeed present such a case,¹¹⁰ it is not clear how generalizable the model is to other settings involving trade in goods embodying intellectual property where the welfare implications of price discrimination are highly ambiguous.¹¹¹ Therefore, if such rules are desirable for pharmaceuticals they could be incorporated into the already existing elaborate system of drug regulation, but there is no need to subordinate the entire area of IP law for that goal.

More generally, recognizing that the welfare effect of price discrimination is ambiguous, contemporary literature believes that there is no compelling basis for antitrust law to prohibit price discrimination.¹¹² It does not follow, however, as proponents of the weak versions of the first sale doctrine suggest, that the law should encourage price discrimination by outlawing arbitrage. Just as it is “difficult for legislators to devise a detailed law that would only attack price discrimination that is injurious to competition and to consumer welfare,”¹¹³ it is difficult to devise laws that would attack only injurious arbitrage.

B. Other Efficiencies of Post-Sale Restraints

1. Parallel trade and pre-sale and post-sale services

In addition to the perceived benefits of price discrimination, antitrust scholarship has identified a variety of other benefits arising from post-sale restraints. The marketing of almost every good requires some level of local investment in establishing a distribution network, complying with local regulations, creating demand through advertisement and other marketing efforts, and offering pre-sale services for potential customers (such as demonstrations, training, and advice) as well as post-sale services (such as technical support, handling repairs and returns, and respecting warranties).¹¹⁴ A

110. It is not clear that they are. *See supra* notes 105–09 and accompanying text.

111. TREBILCOCK ET AL., *supra* note 65, at 346.

112. *Id.* at 371.

113. *Id.* at 351.

114. *See generally* Iacobucci, *supra* note 48 (providing a survey and critique of various

manufacturer that is fully integrated into distribution and retail can ensure, through managerial orders, that all of these tasks are fulfilled. However, full vertical integration may not be possible or even desirable, and the manufacturer may need to enter into contracts with local firms to provide these services.¹¹⁵ Local dealers (distributors or retailers) may be reluctant to make such investments if they are concerned that after incurring these costs consumers will buy from other dealers offering the goods at a lower price. The manufacturer may then impose various vertical restraints, such as territorial restraints (limiting the dealers to sell only within a designated territory) or price restraints (RPM). If the purpose of such restraints is indeed to guarantee investment and those investments are indeed necessary to increase output, then these restraints are efficient because they will result in more units sold and better services delivered.¹¹⁶

From this perspective, parallel trade is a cause for concern because it allows the parallel importer and the foreign dealer to free ride on the investments of the local dealer. Parallel trade thus could undermine the incentive to invest in building the local market and to provide pre-sale and post-sale services, ultimately to the detriment of the local dealer, local consumers, and the manufacturer.

2. Parallel trade and its effect on positional goods

The appeal of certain products lies in the exclusive status that their high prices confer,¹¹⁷ and according to Barak Orbach, RPM is not only one of the techniques that producers of such products employ to preserve their status, but indeed this technique is “socially superior . . . to other production methods for status goods.”¹¹⁸ An effort by a producer of such a status good to maintain *different* high prices in different countries may be undermined if parallel importation is possible and it has been argued that, indeed, the availability of cheaper products sold via unauthorized outlets may

justifications for vertical restraints).

115. Hovenkamp, *supra* note 20, at 489.

116. Iacobucci, *supra* note 48.

117. Barak Y. Orbach, *The Image Theory: RPM and the Allure of High Prices*, 55 ANTITRUST BULL. 277, 279 (2010).

118. *Id.* at 303.

undermine the value or the prestige of brands, especially luxury brands.¹¹⁹

But while this may be a concern for producers of luxury brands, or for consumers who have purchased them in the pursuit of social distinction, it does not necessarily follow that the state should step in and enforce IP laws to remedy their concerns. As Orbach argues, antitrust law should be agnostic to RPM when it is adopted to maintain the image of luxury goods, because whether the phenomenon of status goods is desirable or not, antitrust laws should not be used to regulate their production.¹²⁰

But in line with Orbach's observation, IP law should be equally agnostic to the sale of positional goods. Just as antitrust law should not be used to discourage trade in such goods, IP rules should not be relied on to encourage it.¹²¹

C. The Irrelevance of Antitrust Insights to the Parallel Trade Debate

From price discrimination to brand image, modern antitrust scholarship, shaped by modern economics, has identified a variety of benefits that post-sale restraints promote, and the identification of these benefits has contributed to the gradual erosion of antitrust law's hostility towards vertical restraints. As the preceding sections have shown, these benefits have often been invoked to support bans on parallel imports, to justify adherence to national exhaustion rules, and to reject international exhaustion. This section explains why, notwithstanding the importance of these insights, they have very limited bearing on the question of whether IP law should adopt a model of national, international, or regional exhaustion.

Consider price discrimination. Even if price discrimination is efficient and arbitrage is harmful and should be prevented by law, banning parallel trade supports only one type of price discrimination: cross-country price discrimination. But price discrimination along

119. See, e.g., NAT'L ECON. RESEARCH ASSOCS. ET AL., *supra* note 108, at 93.

120. *Id.*

121. Indeed, as Barton Beebe has observed, some aspects of modern IP law have developed to establish a new "sumptuary code," which he describes as socially and technologically reactionary, and in contrast with IP's progressive side, intended "[t]o promote the Progress of Science and useful Arts." Barton Beebe, *Intellectual Property Law and the Sumptuary Code*, 123 HARV. L. REV. 809, 814 (2010) (quoting U.S. CONST. art. I, § 8, cl. 8).

national lines is a very crude type of price discrimination. Enforcing it vigorously makes sense only when consumer demand within national boundaries is homogenous but demand across nations is heterogeneous. Generally, however, the preferences of consumers within a nation will be just as varied as the preferences of consumers across nations. Therefore, it is unclear why IP law should target and prevent international arbitrage but remain deferential to interstate, intercity, or interpersonal arbitrage, and there is a serious mismatch between the goal and the legal tool that is supposed to achieve it. In fact, if encouraging price discrimination and preventing arbitrage is the goal of IP law, then the only sensible rule is no exhaustion at all.

Consider the ice skates example discussed earlier and add in consumers in Michigan who have the same demand for ice-skates as their neighbors in Ontario. In addition, assume that some ice skates are patented or have a copyrighted logo embedded in them. If price discrimination is desirable, then the optimal pricing would be to set a higher price in Ontario and Michigan and a lower price in Florida, and if IP law were to guarantee this pricing structure, it should prohibit the unauthorized importation of skates from Florida into Michigan just as it should prohibit their importation into Ontario. Note, however, that a national exhaustion rule will prohibit trade only between Florida and Ontario, but not between Florida and Michigan. Consumers in Michigan will be able to import cheaper skates from Florida if the prices in Michigan were too high, but consumers in Ontario will not. National exhaustion will help segregate the market, but on the basis of an economically irrelevant dimension: national borders, not differences in demand.

Moreover, national exhaustion does not support other forms of price discrimination and the elaborate methods that sellers often employ to maintain them. If IP law should intervene to prevent arbitrage, then national exhaustion misses the mark. Not only does it segregate markets according to an economically irrelevant criterion, but it also targets international arbitrage, where higher natural barriers often exist, while neglecting domestic arbitrage, which often faces no such barriers. National exhaustion operates

where it is less necessary.¹²² Therefore, if exhaustion should be limited to facilitate price discrimination, then logic dictates that no exhaustion ought to be the rule, not only national exhaustion.

In sum, while the virtues of price discrimination are often invoked to support national exhaustion rules, there is a serious mismatch between the symptom (arbitrage) and the remedy (national exhaustion). If arbitrage is a problem, and should IP law be harnessed to prevent it, no exhaustion should be the rule. If the goal is preventing arbitrage, it seems odd to prevent only cross-border arbitrage but permit all forms of domestic arbitrage. Alternatively, and preferably, IP law, which has no power to prevent domestic arbitrage, should simply stay out of the picture of international arbitrage. In fact, the ubiquity of price discrimination schemes that occur domestically, notwithstanding the first sale doctrine, should discount the arguments highlighting the necessity to prevent arbitrage.

The same point can be made with regard to most of the other identified benefits of vertical restraints. It is indisputable that establishing and maintaining efficient distribution systems benefits producers and consumers alike, and that, as I note in the previous sections, often this goal requires the imposition of enforceable post-sale restraints. If IP law should be asked to play a role in achieving better efficiency, national exhaustion provides only a very partial tool. For example, a manufacturer might assign exclusive territories to different dealers or appoint one dealer to deal with one type of customer and another dealer with another type. Each of these dealers might need to make specific investments and might be reluctant to do so without being offered credible protection against free riding by other dealers or the manufacturer. Arguably, the first sale doctrine

122. This conclusion does not exclude the possibility that national borders do provide an economically sound basis for distinguishing between domestic and international arbitrage. If, for example, parallel trade is a response to exchange rate fluctuations, then a cross-border arbitrage may be more prevalent than domestic arbitrage (and arguably more harmful to the local distribution system). While there is evidence that parallel importers are sensitive to currency fluctuations, the significance of this single factor is less than clear. *See* ROSE ANN MACGILLIVRAY, *PARALLEL IMPORTATION* 27 (2010). In any event, it is not clear that manufacturers cannot provide sufficient protection to their local dealers against parallel trade by resorting to contractual means or that there are other effective ways to hedge the risks of currency fluctuations.

undermines the credibility of such guarantees because third parties might be able to obtain the goods and undercut the local dealer. If exhaustion rules should be limited to prevent that scenario, national exhaustion seems like an arbitrary and unprincipled choice because it targets only one type of arbitrage, and not necessarily the most significant one. The only principled choices are either no exhaustion at all or, if one does not find the arguments about the detrimental effects of the first sale doctrine compelling, international (or indeed universal) exhaustion.

While the myriad virtues of post-sale restraints provide very little help in settling the debate about exhaustion rules in the context of parallel trade, it might be tempting to conclude that the first sale doctrine, by undermining the efficacy of post-sale restraints, is indeed an inefficient IP rule, a relic from an era in which the economics of vertical restraints were not well understood. However, the next Part suggests that such a conclusion should not be arrived at precipitously. Recognizing that some post-sale restrictions could be efficient and acknowledging that antitrust law moved in the right direction when it stopped treating agreements imposing such restraints as per se unlawful is one thing. But it hardly follows that post-sale restraints should be enforced as a rule of property, let alone a default one.

VI. THE PROPERTY/CONTRACT FLAW

Modern antitrust teaches us that post-sale restraints are not necessarily harmful and that they may actually be quite beneficial and necessary to organize sophisticated distribution systems when a manufacturer is not fully integrated into distribution and retail.¹²³ The following line of reasoning may thus be adopted: if post-sale restrictions are efficient, they should be enforceable, and if enforcing them on the grounds of IP infringement is easier than on the grounds of breach of contract, an IP remedy should be available. Unfortunately, this line of reasoning is flawed. Remedies for

123. *E.g.*, Hovenkamp, *supra* note 20, at 489 (“[V]ertical restrictions are a very important compromise between unrestricted market transactions and vertical integration through ownership. They permit business entities to have some of the advantages of the market-displacing mechanisms of the business firm but without all of the costs that outright ownership entails.”).

infringement of an IP right may be more effective than those available for breach of contract. However, whether greater efficacy is desirable depends not only on the benefits of more compliance, but also on the costs that may be externalized to third parties.¹²⁴ Therefore, before concluding that greater enforceability is better, it is important to carefully understand why and against whom post-sale restrictions may need to be enforced.

Whether post-sale restraints are imposed to facilitate geographic price discrimination, to maintain the status of a luxury good, or to encourage investment in building a distribution system, or in the provision of pre-sale or post-sale services, a closer look at the antitrust scholarship reveals that it has focused on the relationships between collaborating firms attempting to organize an efficient production and distribution system. We can generalize this insight to the following proposition:

Proposition 2: When firms jointly participate in a productive enterprise that is prone to opportunism, various enforceable restrictions may be necessary.

Additional support for Proposition 2 can also be derived from the New Institutional Economics literature that explores the connection between the theory of the firm and intellectual property rights.¹²⁵ Intellectual property rights may be important for overcoming the acute problems arising from the incompleteness of contracts when firms involved in joint production need to contract around information. In these cases, information presents a unique challenge for efficient contracting. Authors or inventors will often lack the complementary assets necessary for successfully commercializing their intellectual goods.¹²⁶ Therefore, they will need to collaborate with coproducers, and they will need to disclose their

124. Henry E. Smith, *Toward an Economic Theory of Property in Information*, in RESEARCH HANDBOOK ON THE ECONOMICS OF PROPERTY LAW 104 (Kenneth Ayotte & Henry E. Smith eds., 2011), available at <http://ssrn.com/abstract=1712089>.

125. See generally Dan L. Burk, *Intellectual Property and the Firm*, 71 U. CHI. L. REV. 3 (2004).

126. See, e.g., David J. Teece, *Capturing Value from Technological Innovation: Integration, Strategic Partnering, and Licensing Decisions*, 18 INTERFACES 46 (1988).

new information for this purpose. However, after *A* (author/inventor) has disclosed her new information to *B* (publisher/producer), the information cannot be untaught. This creates an acute problem. *A* knows that *B*, having learned the new information, may renege on whatever promises were given and therefore will demand to be paid in full in advance. *B*, not knowing what the information is and whether it is valuable at all, will not agree to pay until after learning what the information is.¹²⁷ Even if the parties can ultimately draft a suitable contract, the contract will be inherently incomplete because the information under question cannot be easily defined. Further, any contract will not be binding on third parties who may have learned the information from *B*. IP rights facilitate knowledge-transfer and coproduction by creating an enforceable asset that one can contract about. This reduces the cost of “detering opportunistic rent seeking by team members while simultaneously increasing trust and cooperation within the team.”¹²⁸

Proposition 2 is consistent with Stephen Maurer and Suzanne Scotchmer’s “profit neutrality” principle, which they define as “the principle that a patentholder’s reward should not depend on whether he has the ability to work the patent efficiently himself.”¹²⁹ Applied to the present context, “profit neutrality” means that IP owners should be able to impose post-sale restraints to achieve the same profit that they would be able to gain if they were fully integrated into distribution and retail. In such situations of joint production, treating IP rights as default entitlements and permitting parties to contract around them to achieve efficiency in a profit-neutral way makes sense.¹³⁰ However, the teachings of the literature that supports Proposition 2 cannot immediately apply to restraints that *do*

127. If the information is tacit and *B* will not learn it unless *A* spent the time and effort to teach it, *B* will be concerned that after receiving the payment, *A* may fail to perform its obligation to teach. See Ashish Arora, *Licensing Tacit Knowledge: Intellectual Property Rights and the Market for Know-How*, 4 *ECON. INNOVATION & NEW TECH.* 41, 43 (1995).

128. Paul J. Heald, *A Transaction Cost Theory of Patent Law*, 66 *OHIO ST. L.J.* 473, 489 (2005).

129. Stephen M. Maurer & Suzanne Scotchmer, *Profit Neutrality in Licensing: The Boundary Between Antitrust Law and Patent Law*, 8 *AM. L. ECON. REV.* 476, 480 (2006).

130. See also Anne S. Layne-Farrar, *An Economic Defense of Flexibility in IPR Licensing: Contracting Around “First Sale” in Multilevel Production Settings*, 51 *SANTA CLARA L. REV.* 1149, 1184–85 (2011).

not involve situations of joint production and cannot generally support the imposition of enforceable post-sale restraints against third parties.

For example, even if and when territorial restrictions should justifiably protect a local dealer from opportunistic behavior by other members of the same distribution system (other dealers or the manufacturer), it does not directly follow that third parties who are not members of the same distribution system and who lawfully purchased a product in another territory should be prohibited from reselling them in the local territory. The cost/benefit calculus of making the territorial restriction enforceable against third parties is not the same as the calculus of enforcing the restraints against members of the distribution system. Although restraining the third party might be efficient—inasmuch as its activities and interests are not fully aligned with those of members of the distribution system—the likelihood that sales made by third parties will seriously undermine the viability of the authorized distribution system diminishes with distance from the authorized channel, as well as over time. Therefore, the need to control third parties diminishes too. At the same time, if restraints were to be enforceable against third parties whose identity cannot be determined *ex ante*, they should come in the form of an *in rem*, or property, entitlement. The trouble is that if restraints are enforced *in rem*, they may be enforced even when and where the reasons for imposing them no longer exist, or in situations where enforcing them may not be socially beneficial. Moreover, such an *in rem* right might interfere without good reason with buyers' freedom to seek the best deal available, freedom which drives competitive markets and innovation in the first place, or with buyers' freedom to use the goods that they purchase in the best way they see fit, including as inputs in their own innovative endeavors.

For example, consider a company introducing an innovative product such as the iPad. The benefits of post-sale restraints are easy to describe. Assume that successfully marketing such a product requires setting up a distribution network that involves at least some retailers who can display the new product, allow potential consumers to touch and feel it, and provide on-site demonstrations and training. Retailers providing such services face higher costs than retailers who do not, and they might be reluctant to provide them if consumers could take advantage of these services but then purchase the product from another retailer (e.g., online) selling the product at

a discount. This is the paradigmatic case in which post-sale restraints, limiting the ability of one retailer from undercutting the other, might be justified. Such post-sale restraints could be imposed by contract and if the product is desirable enough for the retailer, the threat of termination might be enough to secure compliance. However, contractual restraints are enforceable only against the contractual parties and would not protect the retailer from competition from other sellers who are not bound by any contractual restriction.

Or, for example, suppose that the manufacturer has implemented a price discrimination scheme under which some dedicated retailers sell the product to some consumers at a significant discount.¹³¹ Such retailers might be bound contractually from competing with the retailers in the other segment, but these contracts will not bind third parties, such as arbitrageurs buying the product from these dedicated retailers at the discounted price and then reselling it in competition with the ordinary retailers. This outcome could be averted if the manufacturer were permitted to opt-out of the first sale doctrine, especially since a product such as an iPad comprises numerous patented components and is loaded with copyrighted software. If the first sale of the iPad does not exhaust these patents and copyrights, arbitrage could be prevented more easily because every unauthorized reseller would infringe those rights.

So far so good, but there is a downside too. If the sale of the iPad does not exhaust these patents and copyrights, any further transfer of the iPad would require the permission of the relevant patent and copyright owners (which may or may not be the producer) for the entire duration of the patents or the copyrights. Consequently, the buyer would not be permitted to sell her old iPad a few (or even many) years down the road, not even when the product becomes obsolete, without violating IP rights. It is evident that such long-lasting restrictions, which apply in rem, are not necessary for solving the original problem, namely opportunism among retailers. Further, the restrictions will remain valid even when opportunism is no longer a concern, and the duration of such

131. For example, consider Apple's education pricing. See *Apple Store for Education*, APPLE, <http://store.apple.com/us-hed/findyourschool> (last visited Jan. 13, 2014).

restrictions and their in rem applicability may impose diverse costs on future buyers and users, costs that the producer and the first buyer will not internalize because they may be borne by third parties, or in any event will be heavily discounted because they may be remote or unknown.

Therefore, while the availability of remedies for IP infringement reduces the costs of enforcing post-sale restrictions, such remedies may increase the social costs associated with such restraints. This leads us to the next proposition:

Proposition 3: The economics of joint production do not normally justify long-lasting post-sale restraints or restraints imposed on third parties. Therefore, enforcing post-sale restraints should not normally be part of the property bundle.

Proposition 3 is crucial to the design of an optimal exhaustion rule because, if true, it would reject the weaker forms of exhaustion. It is important to elaborate not only on why and when post-sale restraints are efficient, but also on the optimal way to enforce them is. In other words, formulating the optimal exhaustion rule requires not only an understanding of when post-sale restraints are efficient, but also an ability to determine whether they should be enforced as a matter of contract law (and perhaps tort law) or IP law, and against whom they should be enforced.

A. The Costs and Benefits of Infringement Remedies

In his recent analysis of the first sale doctrine, Herbert Hovenkamp embraced the superiority of enforcement through infringement actions over suits for breach of contract and criticized the U.S. government for taking the position that post-sale restrictions should be enforced solely as a matter of contract law, for which it argued in its amicus brief in *Quanta*.¹³² Hovenkamp criticized the government (as well as the Court for deciding the case in a manner consistent with the government's position) for being

132. Brief for the United States as Amicus Curiae Supporting Petitioners, *Quanta Computer, Inc. v. LG Elecs., Inc.*, 553 U.S. 617 (2008) (No. 06-937), 2007 WL 3353102, at *5.

more concerned about form than substance,¹³³ and as “excessively draconian”¹³⁴ for yielding a per se rule that denies automatically “with no consideration of the restraint’s purpose or effect”¹³⁵ any IP remedy once a sale is found. Hovenkamp lists several advantages of allowing IP remedies for breach of post-sale restraints: their applicability to a large and diverse number of downstream people;¹³⁶ availability of injunctive relief;¹³⁷ the calculation of damages on the basis of a reasonable royalty together with interest and costs (and up to treble damages for willful patent infringement) rather than expectation damages;¹³⁸ and a possible jurisdictional advantage of suing in a federal rather than in a state court.¹³⁹ In addition to these legal advantages, Hovenkamp emphasizes that although theoretically a seller might be able to bind all downstream purchasers contractually by demanding that the initial purchaser not only agree to the restriction, but also, upon pain of damages, impose the restriction on all parties downstream, this solution would be messy and would increase transaction costs significantly.¹⁴⁰ Like a servitude on real property that provides a neater and cheaper solution, enforcing a post-sale restriction as a matter of IP law would be efficient at least as long as a purchaser has timely notice of the restriction.¹⁴¹

Hovenkamp is not the first commentator to parallel the common law’s tolerance towards land servitudes with the first sale doctrine’s relationship to post-sale restraints.¹⁴² However, upon closer inspection, rather than confirming the superiority of enforcing post-sale restraints with IP remedies, the analogy to land servitudes demonstrates the inferiority of such products.

133. Hovenkamp, *supra* note 200, at 540–41.

134. *Id.* at 546.

135. *Id.* at 541.

136. *Id.*

137. *Id.* at 543.

138. *Id.*

139. *Id.* at 544.

140. *Id.* at 541.

141. *Id.* at 541–43, 546.

142. *See, e.g.*, Glen O. Robinson, *Personal Property Servitudes*, 71 U. CHI. L. REV. 1449 (2004); *see also* Molly Shaffer Van Houweling, *The New Servitudes*, 96 GEO. L.J. 885 (2008) (presenting a critical survey of this literature).

The economics of efficient land use and the economics of efficient distribution (and use) of goods embedding IP differ substantially. Although both land servitudes and post-sale restraints may solve organizational problems and foster more efficient asset use, as I explain below, the organizational problems are different, and the negative effects of enforceable servitudes/restraints are different. Consequently, the costs and benefits of various instruments for enforcing the restraints (e.g., contract versus property rules) are not the same, and it would be an unfortunate error to apply uncritically lessons from one area to the other.

Despite early skepticism, the common law has grown to accept the utility of land servitudes as a means of private ordering to coordinate productive land use and reduce potential conflicts in ways that supplement and sometimes substitute for land-planning law.¹⁴³ Increased urban density¹⁴⁴ and the inevitable externalities that the use of one piece of land creates for others, coupled with the finite nature of land, render indispensable the need for restrictions on use. Further, the attributes of land suggest that the need for coordination is as durable as the land itself, is unlikely to decrease over time, and will often outlive the tenure of those who own or possess the land at the time the need to coordinate arises. Under such circumstances, the cost-inefficacy of reliance on contracts to enforce land-use restrictions is evident and the superiority of servitudes is clear.

The common law's early reluctance towards servitudes has traditionally focused on three types of concerns: the problem of notice and information costs that may arise when purchasers of burdened land may not be aware of restrictions limiting the use of the purchased asset;¹⁴⁵ the problem of limitations on the freedom of future generations to manage resources wisely, autonomously, and efficiently;¹⁴⁶ and concerns over other externalities, such as restraints on trade and competition or the limitation on individuals' fundamental rights.¹⁴⁷ However, the development of registration

143. Van Houweling, *supra* note 142, at 892.

144. *Id.*

145. *Id.* at 893.

146. *Id.* at 900.

147. *Id.* at 905.

systems (together with notice-encouraging doctrines)¹⁴⁸ and the fact that the value involved in land transactions typically justify incurring the costs of learning about potential land restrictions, have minimized the notice problem associated with land servitudes. The problem of the future may still be of concern in land, but it is precisely the need to coordinate land use over long periods of time that creates the need for enforceable use-restrictions and highlights the superiority of enforcing restrictions as servitudes rather than by contract. It makes sense to prefer solutions for known and pressing needs arising today that affect the foreseeable future and to discount the possibility that such solutions would create possible but undeterminable problems in the distant future. Similarly, the need to impose restraints to internalize external costs weighs in favor of land servitudes, while the local nature of land makes it less likely that land servitudes will systematically result in a significant anticompetitive harm or otherwise violate important public policies. When these negative consequences do occur, antitrust law, the common law of restraints on trade,¹⁴⁹ and the ability of courts to invalidate such servitudes¹⁵⁰ may effectively weed out the harmful restraints.

From an organizational perspective, both land servitudes and post-sale restraints may promote productive asset use in situations when integration is incomplete or absent. Obviously, if one firm owned all tracts of land in a relevant area, it would internalize all costs associated with land use and would not need to create servitudes because it could decide for itself the optimal way to use the land. Likewise, a firm that is fully integrated into distribution and retail would not need to impose post-sale restraints because it could control all externalities between its units through managerial orders. Beyond these similarities, however, land servitudes and IP servitudes (that is, post-sale restraints enforceable as a matter of IP law), are mirror images of each other.

Unlike land, whose efficient use may require durable restraints that would be enforceable against future buyers, the organizational

148. See generally Herbert Hovenkamp, *Notice and Patent Remedies*, 88 TEX. L. REV. 221 (2011).

149. See, e.g., MICHAEL TREBILCOCK, *THE COMMON LAW OF RESTRAINT OF TRADE: A LEGAL AND ECONOMIC ANALYSIS* 307 (1986).

150. Van Houweling, *supra* note 142, at 905.

problems in the context of collaboration in the production and distribution of innovative goods occur mainly at the early stages of the product life-cycle:¹⁵¹ production problems disappear immediately after the good is produced,¹⁵² and distribution problems largely disappear upon the distribution of the good or shortly thereafter. Notably, these problems cease to exist long before the IP rights expire (particularly in the case of copyright). Moreover, as coproducers, the firms are in privity and this enables them to rely on contracts for addressing many of the organizational problems associated with efficient distribution.¹⁵³ Further, because the standard remedy for breach of contract is only damages, and the plaintiff needs to prove actual damage, a party bound by a contractually valid post-sale restraint may be able to put the good to better use when no damage can be shown or otherwise when the breach is efficient. Thus, while IP remedies may increase the enforceability of post-sale restraints, the benefit is marginal and must be weighed against the fact that the threat of IP remedies might prevent the efficient use of the good long into the future, even when they no longer serve their original purpose.

151. Ariel Katz, *Substitution and Schumpeterian Effects over the Life Cycle of Copyrighted Works*, 49 JURIMETRICS J. 113 (2009) [hereinafter Katz, *Substitution*].

152. *Id.*

153. As opposed to organizational problems during the development and production stages, which contract law, alone, may not solve, and which IP rights may help to ameliorate. *Id.* at 141–42.

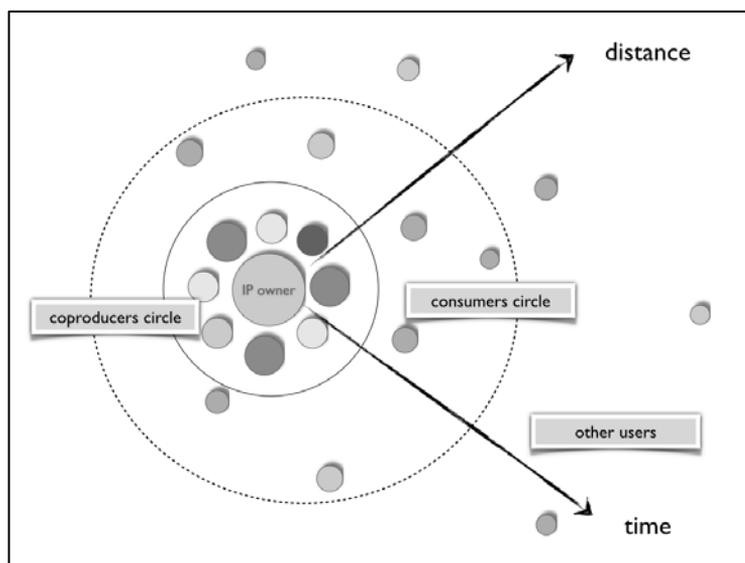


Figure 1 An IP owner, coproducers, and users

Figure 1 is a graphic representation of the circles of relationship between an IP owner and interested parties through the life-cycle of an innovative good. The inner circle is the IP owner and its coproducers, or the several firms that participate in the production or initial distribution of an innovative good. In the second circle are the consumers who buy the goods produced in the inner circle. The third circle is populated by the users who may at one point be interested in obtaining and using the innovative good but do not obtain it directly from the inner circle. Two vectors—distance and time—are also presented. Distance reflects the transactional proximity between the IP owner and the user. As the distance between the user and the IP owner increases, concerns about opportunism diminish. The time vector reflects temporal proximity between production, distribution, and use.

As noted above, the benefits of post-sale restraints are concentrated primarily within the inner circle of the IP owner and the several firms that participate in the production or initial distribution of an innovative good. Hence, the marginal benefit from having enforceable restraints diminishes as we move along the vectors of distance and time. At the same time, the marginal social costs associated with goods encumbered by restraints could easily increase over distance and time because any use inconsistent with the

restraint would require the IP owners' permission, yet the cost of obtaining such permission could easily increase over time and distance. I will discuss the reasons why we should care about these costs in Part VII.

In sum, even though post-sale restraints might be efficient and justifiably enforceable, it does not follow that enforcing them should be part of the bundle of property rights available to IP owners, allowing them to reach out in distance and time and impose long lasting restraints of diminishing social utility. The first sale doctrine, therefore, reflects the law's sensitivity to the differences between the costs and benefits of contract as opposed to property entitlements. The law grants strong property rights that can promote efficient contracting between coproducers at the early and critical stages of production and distribution, but guarantees that trade will not be encumbered through durable, but often unnecessary, restraints later on.

VII. THE (NON)-COASIAN FLAW

It follows from the preceding discussion that when post-sale restraints are efficient, they should generally be imposed and enforced as a matter of contract law, not property. However, there might be exceptions justifying a resort to IP remedies. For example, if it could be demonstrated that a life-saving drug would not be produced or would not be sold in developing countries unless all opportunities of arbitrage are eliminated, and that IP remedies are necessary, then we should hesitate ruling them out. It is also possible that the use of "viral" restrictive conditions common in open-source software licenses helps solve long-term coordination problems between noncontracting contributors, making open-source projects similar to land and justifying greater tolerance towards this kind of "new servitudes."¹⁵⁴ Moreover, it is also possible to minimize the potential costs resulting from IP-based

154. See Van Houweling, *supra* note 142; Massimo D'Antoni & Maria Alessandra Rossi, *Appropriability and Incentives with Complementary Innovations* (Dep't of Econ., Univ. of Siena, Working Paper Series, No. 603 (Nov. 2010)), available at <http://www.econ-pol.unisi.it/quaderni/603.pdf>.

restraints through liberal application of various defenses.¹⁵⁵ Arguably, fair use or other defenses (such as misuse, or the repair and construction defense in patent law) can help ensure that inefficient post-sale restraints will not be enforced.

Therefore, the issue is not so much about a categorical choice between IP remedies and contract remedies (and whether clever drafting can guarantee that a restriction is found to be a license condition rather than contractual covenant),¹⁵⁶ but about the design of default exhaustion rules. In other words, the important question is, assuming exhaustion is the default, how strong or sticky this default rule should be. Under what conditions should courts enforce deviations from the default and what are the grounds for invalidating such deviations? Equally important is the question of who bears the burden of upholding or invalidating the restraint and how heavy this burden is. Should the restraint be considered presumptively valid unless proven otherwise, or should contracting around the first sale doctrine be regarded as presumptively invalid unless the IP owner demonstrates its efficiency? In this Part, I advocate the latter option and offer the following proposition:

Proposition 4: Contracting around the first sale doctrine should be presumptively invalid. Courts should refuse to enforce license conditions or contract terms limiting the ability of the user to resell goods embodying IP rights unless the IP owner can demonstrate that the restraint is necessary and superior to other means to achieve efficiency.

To support Proposition 4, I need first to reject the view that exhaustion is no more than a simple default rule that can be easily modified by contracts that would generally be held valid. Current case law tends to reflect this paradigm and as a result, many cases are decided on the basis of relatively marginal legal questions such as what constitutes a valid contract,¹⁵⁷ whether a first conditional sale

155. *Kirtsaeng v. John Wiley & Sons, Inc.*, 133 S. Ct. 1351, 1389 (2013) (Ginsburg, J., dissenting) (arguing that principles of fair use and implied license can ameliorate many of the negative consequences that the majority was concerned about when it favored international exhaustion over national exhaustion).

156. *MDY Indus. v. Blizzard Entm't, Inc.*, 629 F.3d 928 (9th Cir. 2010).

157. *See, e.g., ProCD v. Zeidenberg*, 86 F.3d 1447 (7th Cir. 1996) (discussing the

preempts exhaustion,¹⁵⁸ whether there was a sale or just a license,¹⁵⁹ and whether restraints imposed by notice are sufficient.¹⁶⁰ These cases seem to be premised on the view that the first sale doctrine is merely a simple default rule, which can be worked around as long as the IP owner has used the proper legal tool.

*ProCD v. Zeidenberg*¹⁶¹ represents the classic articulation of this paradigm. While strictly speaking, *ProCD* is not a first sale case, it is relevant because it stands for the proposition that IP rights create only default entitlements from which transacting parties can freely deviate by contract and that such contract terms are generally enforceable “unless their terms are objectionable on grounds applicable to contracts in general (for example, if they violate a rule of positive law, or if they are unconscionable).”¹⁶² In this decision, Judge Easterbrook was clear that deviating from the entitlements set by copyright law does not “violate a rule of positive law” because the public policy embedded into copyright law’s choice of entitlement cannot be affected by contracts.¹⁶³ In other words, property is property and contracts are contracts. Unlike a copyright, which is “a right against the world,”¹⁶⁴ “[c]ontracts generally affect only their parties, [and] do not create ‘exclusive rights.’”¹⁶⁵

According to Judge Easterbrook, license terms accompanying works or products are conceptually identical to any other feature of the products. Both determine what the consumer can do with the product, and both are reflected in the product’s price and mediated through the market. Generally, courts do not design products, do not determine their prices, and therefore should not intervene in the

validity of shrink-wrap licenses and holding that they are valid, as long as the buyer can return the product after having an opportunity to read the terms).

158. *Quanta Computer, Inc. v. LG Elecs., Inc.*, 553 U.S. 617 (2008) (leaving the question open).

159. *Vernor v. Autodesk, Inc.*, 621 F.3d 1102 (2010).

160. *UMG Recordings, Inc. v. Augusto*, 628 F.3d 1175, 1180 (9th Cir. 2011) (distinguishing *Vernor* on—among other things—lack of acceptance of the restrictions).

161. 86 F.3d 1447 (7th Cir. 1996).

162. *Id.* at 1449.

163. *Id.*

164. *Id.* at 1454.

165. *Id.*

conditions accompanying their sale: “Competition, not judicial oversight, is the best protector of consumers’ interests.”¹⁶⁶

The logic that underlies the *ProCD* judgment arises from the Coase Theorem that asserts that in the absence of transaction costs, bargaining will lead to an efficient outcome regardless of the initial allocation of property rights.¹⁶⁷ Assuming away external effects, or as the judgment stated, “[c]ontracts generally affect only their parties,” the distinction between property and contract provides a seemingly elegant answer to the question of whether there is any problem in contracting around limitations to IP rights. Assuming that such restrictions affect only their parties, a contract suggests that—save for some extraordinary circumstances—the transaction must be efficient. Assuming also that markets are perfectly competitive and consumers are fully informed, license terms can indeed be treated like any other product feature. Indeed, in a lecture given shortly before he rendered his decision in *ProCD*, Judge Easterbrook explicitly referenced the Coase Theorem when he stated that given the difficulty of setting optimal IP rules, the most sensible approach is to encourage Coasian bargaining by creating clear and enforceable property rights, facilitating bargaining, and “enjoy[ing] the benefits.”¹⁶⁸

In a Coasian world, it does not matter whether the first sale doctrine exists or not because transacting parties will always be able to efficiently bargain about the rights to resale or otherwise use an item. If resale is efficient, the owner and the user will enter into a contract permitting it, and if it is not, the contract will restrict it regardless of the initial allocation of the resale right. If some consumers value the ability to resell the good more than others, then producers would be happy to sell the goods with or without such rights at different prices. Under the logic used in the *ProCD* case, we should not worry about the question of exhaustion. All that the law has to do is treat the first sale doctrine as no more than a baseline for contracting, permit and uphold all subsequent bargaining, and then let us all “enjoy the benefits.”

166. *Id.* at 1453.

167. Ronald Coase, *The Problem of Social Cost*, 3 J.L. & ECON. 1 (1960).

168. Frank H. Easterbrook, *Cyberspace and the Law of the Horse*, 1996 U. CHI. LEGAL F. 207.

While attractive, this policy prescription is flawed. It is flawed because the world of IP is notoriously non-Coasian.¹⁶⁹ The need for intellectual property rights, and the need to define their limits, arises precisely because the world of information is non-Coasian.¹⁷⁰

In a Coasian world, there would be no need for IP rights because inventors or authors could secure the necessary return on their investment by contracting with potential users.¹⁷¹ They would pitch their ideas to potential users and investors, and interested parties would undertake to finance it. Unfortunately, relying solely on contracts may lead to suboptimal levels of investment due to several market failures. For example, as Kenneth Arrow has observed, an investor may be reluctant to invest in another person's innovation before learning what the innovation is, but will have no reason to invest once the information has been disclosed and can be used by anyone for free.¹⁷² Other failures may be a result of high transaction costs, such as the need to transact with numerous potential investors; the need to finance before it is clear what, if anything, will ultimately be invented or created and how valuable will it be; and the realization that investors might be better off waiting until the good is created through the investment of other parties and then copy it for free. Intellectual property rights are needed precisely for addressing these and similar market failures.

One may concede that even if prior to creating the intellectual good the world is non-Coasian, Coasian bargaining over the rights to use such goods, once created, is entirely possible. Arguably, then,

169. The term "Coasian" is misleading because it actually ignored Coase's main contribution, namely, that transaction costs are pervasive and important and that the law does matter. As Coase himself wrote, "[t]he world of zero transaction costs has often been described as a Coasian world. Nothing could be further from the truth. It is the world of modern economic theory, one which I was hoping to persuade economists to leave." See RONALD H. COASE, *THE FIRM, THE MARKET, AND THE LAW* 174 (1988).

170. Indeed, this is true for any type of property, not only IP. See Thomas W. Merrill & Henry E. Smith, *Making Coasian Property More Coasian*, 54 J. L. & ECON. 577, 93 (2011) ("We have property and endow it with a basic architecture of exclusion rules supplemented by rules and standards governing proper use, precisely because of transaction costs.").

171. See Harold Demsetz, *Information and Efficiency: Another Viewpoint*, 12 J. L. & ECON. 1, 12 (1969). See also Matthew Sag, *Beyond Abstraction: The Law and Economics of Copyright Scope and Doctrinal Efficiency*, 81 TUL. L. REV. 187, 209 (2006).

172. KENNETH ARROW, *ESSAYS IN THE THEORY OF RISK-BEARING* 152 (1976); HENRY CHESBROUGH, *OPEN BUSINESS MODELS* 68 (2006).

once the law grants an IP right, any subsequent bargaining with respect to its use should be regarded as presumptively efficient. This view seems to be implied in Judge Easterbrook's position.¹⁷³ The problem with this view is that if the post-creation world were indeed Coasian, there would be no need to limit the duration and scope of IP rights;¹⁷⁴ the term of exclusivity would be mediated through the market and the price system. There would be no need for fair use or other limitations on owners' rights because the socially optimal outcome would be achieved even if owners had total control. Any socially efficient use of the work would be authorized: "You want to criticize my work? Go ahead, here's a license"; "You want to build on my ideas to develop a better product that will displace my own? No problem, here's your license." In a perfectly Coasian world, what IP law allocates to owners and users respectively would serve only as a starting point from which transacting parties would freely negotiate to maximize their own, and society's, gains.¹⁷⁵ Indeed, such allocation could be totally arbitrary as long as the law clearly defined what was allocated to whom and then enforced all subsequent contracts.

But our world is not perfectly Coasian, and limitations on IP rights play an important role in the real world.¹⁷⁶ Therefore, we cannot assume that IP law provides only the baseline from which bargaining will necessarily, or even presumptively, increase social welfare.¹⁷⁷ This does not mean that contracting out of limitations on IP rights cannot increase social welfare; we have no reason to assume that the initial allocation is always optimal, and we have grounds to believe that sometimes it may not be. However, if the law limits the

173. Easterbrook, *supra* note 168.

174. Cf. Merrill & Smith, *supra* note 170, at 32 ("In a world of zero transaction costs, it would not matter whether property rights are broad or narrow, clear or ambiguous—or in rem or in personam.").

175. Easterbrook, *supra* note 168.

176. Stephen Breyer, *The Uneasy Case for Copyright: A Study of Copyright in Books, Photocopies, and Computer Programs*, 84 HARV. L. REV. 281, 316 (1970); see also Ariel Katz, *Copyright and Competition Policy*, in HANDBOOK OF THE DIGITAL CREATIVE ECONOMY 209, 215 (Ruth Towse & Christian Handke eds., 2013) [hereinafter Katz, *Copyright and Competition*].

177. Except, perhaps, in the case of questions of who is the first owner, as opposed to the question of what this ownership entails. IP law clearly contemplates assignments of ownership, or the grant of licenses, which, by definition presuppose the possibility that the owner is not necessarily the best exploiter.

term and scope of IP rights under the assumption that market failures exist, and that transacting parties' private costs and benefits are not fully aligned with social costs and benefits, then it logically follows that there can be no presumption that contracting around such limitation increases social welfare.¹⁷⁸

This is particularly true when intellectual goods are distributed on a mass scale subject to restrictive terms. If it becomes difficult to access intellectual goods to which such terms do not apply, the practical result may be quite similar to an *in rem* obligation. Moreover, when technological measures reinforce the restrictions and the law prohibits circumventing them, the distinction collapses altogether because the restrictions affect everyone in touch with the work—even those who are not privy to the initial contract or the license.¹⁷⁹

Returning to the *ProCD* case, Judge Easterbrook's characterization of license terms as a product feature generally immune to judicial oversight does not support the weight of the *laissez faire* conclusion he arrives at. While it is true that as a general matter competition protects consumers' interests better than judicial oversight, there are at least three reasons why market competition may not create conditions that justify a presumption of efficiency for contracting out of IP limitations.

Firstly, markets for copyrighted and patented works are not perfectly competitive because limiting competition is the *raison d'être* of copyrights and patents.¹⁸⁰ Only the owner or her licensees can be the legal source of copies of any particular intellectual good. There cannot be unfettered competition between different sellers competing over price and license terms. Although competition from other non-infringing works may still exist, assuming that this competition resembles conditions in markets with near-perfect

178. Compare RADIN, *supra* note 109, at 172 (arguing that widespread boilerplate schemes in which fair use and other user rights are contracted around should be scrutinized carefully rather than assumed to be efficient); see also Sag, *supra* note 171, at 212 (“[R]elying exclusively on market mechanisms of exchange creates the danger that strategic exploitation of the market system can reduce aggregate welfare.”).

179. RADIN, *supra* note 109, at 169; see also Katz, *Substitution*, *supra* note 151, at 144.

180. Katz, *Making Sense*, *supra* note 51, at 851.

competition is conceptually flawed and empirically incorrect; if it were true, IP rights would be totally useless.¹⁸¹

Secondly, even if license terms can be conceived as equivalent to product features, the more complex the combination of features/terms/prices, the more likely it is that consumers will not be capable to fully comprehend on what they are contracting.¹⁸² If such information gaps and asymmetries exist, the assumption that the market functions efficiently becomes less credible.¹⁸³ While it is true that competitors might be interested in bridging the information gap in order to increase their own sales at the expense of their rivals, competitors may not always find it worth their while to bridge the information gap,¹⁸⁴ and as noted above, IP law intentionally reduces the intensity of competition anyway.¹⁸⁵

Thirdly, even when competition exists and users are fully informed about the rights they waive, such transactions may not be efficient because the transacting parties will fail to consider the externalities imposed on third parties.¹⁸⁶ Many users' rights (including copyright subject matter limitations) permit users to engage in creative activities that benefit not only the users themselves, but also third parties and society at large.¹⁸⁷ These users,

181. *Id.* at 852.

182. RADIN, *supra* note 109, at 103.

183. *Id.* at 107–08.

184. David Gilo & Ariel Porat, *Viewing Unconscionability Through a Market Lens*, 52 WM. & MARY L. REV. 133 (2010).

185. It should be mentioned that trademark law may also reduce competitors' ability to inform their consumers about the advantages of their products or services over those of their rivals. See Ariel Katz, *Beyond Search Costs: The Linguistic and Trust Functions of Trademarks*, 2010 BYU L. REV. 1555 (2010).

186. Katz, *Copyright and Competition*, *supra* note 176, at 214.

187. *Feist Publ'ns, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340 (1991); *CCH Canadian Ltd. v. Law Soc'y of Upper Can.*, [2004] 1 S.C.R. 339, para. 23 (Can.) (recognizing "society's interest in maintaining a robust public domain that could help foster future creative innovation"); see also *Théberge v. Galerie d'Art du Petit Champlain Inc.*, [2002] 2 S.C.R. 336, para. 32 (Can.) ("Excessive control by holders of copyrights and other forms of intellectual property may unduly limit the ability of the public domain to incorporate and embellish creative innovation in the long-term interests of society as a whole, or create practical obstacles to proper utilization. This is reflected in the exceptions to copyright infringement enumerated in ss. 29 to 32.2, which seek to protect the public domain in traditional ways such as fair dealing for the purpose of criticism or review and to add new protections to reflect new technology, such as limited computer program reproduction and 'ephemeral recordings' in connection with live performances.").

however, can expect to internalize only part of the social benefits arising from their activities and will not take into account the positive spillovers conferred on others.¹⁸⁸ For example, a reader may buy a book and then decide to write a critique of it, or the book may inspire her to write additional works. If that reader has to pay for the right to criticize the book or for the right to be inspired by it, her willingness to pay will reflect only the private value that she might expect to derive from these activities, but not the value that her activities will generate to others. Copyright owners ignore those positive spillovers as well and would be especially reluctant to permit uses that might harm their own interests,¹⁸⁹ or they might strategically exploit situations of hold-up to extract the highest licensing fees possible.¹⁹⁰ Consequently, under conditions of full alienability of user rights, the market will fail to generate socially optimal transactions between owners and users.¹⁹¹

Lastly, even if first innovators and follow-on innovators could internalize all such spillovers, they may still not be able to negotiate effectively because of the same information gaps that prevent efficient contracting between investors and the earlier innovators discussed above.¹⁹² Thus, if we cannot assume that conditions for Coasian bargaining over first and follow-on innovations exist, there is no reason to assume that contracts restricting such follow-on innovations are efficient.

The previous Part challenged the wisdom of making IP remedies generally available for enforcing otherwise valid restraints, noting that IP owners and their transacting parties may not internalize the social costs arising from the long term and in rem nature of IP entitlements. It also suggested that enforcing post-sale restraints should generally be within the domain of contracts. This Part went further and warned against adopting a false Coasian view, namely

188. Brett M. Frischmann & Mark A. Lemley, *Spillovers*, 107 COLUM. L. REV. 257, 262 (2007).

189. *See, e.g.*, *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 592 (1995) (recognizing that “the unlikelihood that creators of imaginative works will license critical reviews or lampoons of their own productions removes such uses from the very notion of a potential licensing market”).

190. *Sag*, *supra* note 171, at 212 (discussing strategic behavior).

191. RADIN, *supra* note 109, at 171.

192. *See supra* notes 170–75 and accompanying text.

that contracting around IP limitations is presumptively efficient socially. The following Part ties together these two insights, specifically in the context of exhaustion and offers a justification for the continuing vitality of the first sale doctrine.

VIII. JUSTIFYING THE FIRST SALE DOCTRINE

The need for (socially efficient) post-sale restraints arises primarily in settings of joint-production and incomplete vertical integration. In such cases, working around exhaustion rules is not objectionable, at least to the extent that contracting around the rules guarantees or improves productive activity (i.e., developing new products, producing, and distributing them). In contrast, restraints in settings that do not involve joint-production (e.g., those imposed on end users or those affecting third parties) result in a different cost-benefit calculus. While relying on IP remedies for enforcing restraints, which makes such remedies available against end users or third parties, might increase the efficacy of the restraint at the margin (e.g., contracts may prohibit one distributor to sell in a territory of another distributor, but the contract will not prevent a third party from buying in one territory and selling in the other), the social costs of IP remedies might outweigh this marginal benefit. The reason is that post-sale restraints that can be enforced in rem continue to exist even when they are no longer justified or would not have been justified in the first place.¹⁹³ Any use that is inconsistent with the restraint would require the copyright owner's permission, but, as I have discussed in the previous Part, even if transaction costs would not render the pursuit of permission prohibitively costly, information asymmetries, strategic behavior, and externalities, all of which are endemic in contracting around intellectual goods, suggest that there is no reason to presume that contracting around IP limitations will be socially efficient. In this Part, I focus on justifying the first sale doctrine and argue that it should be treated as a sticky default rule. I show that the first sale doctrine improves welfare by reducing transaction costs, and I explain why even giving full notice of the restraint does not solve the problem, since IP owners and their transacting parties will not internalize the full social costs of their

193. See *supra*, Part VI.A.

transactions and ignore the full benefits that exhaustion permits.

Proponents of a weak first sale doctrine recognize that at times, post-sale restraints could be used anticompetitively, or cause harm to innovation, but argue that in such cases antitrust law or the doctrine of IP misuse would be sufficient to invalidate them.¹⁹⁴ However, there are additional types of social costs, both static and dynamic, that IP owners and their transacting parties might ignore when they want to maximize their respective private benefits, and that antitrust law is not designed to address and IP misuse is incapable of remedying. The static costs relate to the preservation and resource waste of physical objects, and the dynamic costs involve loss of knowledge and negative impacts on future innovation, particularly user-innovation.

A. Preservation and Resource Waste

Copyright owners discontinue a large number of books and recordings each year. Of the more than ten thousand books published in the United States in 1930, only 174 were still in print in 2001.¹⁹⁵ In 1999 alone, Barnes and Noble stated that ninety-thousand books went out of print. Many of these books are shelved in public libraries and private domiciles, but few remain in publishers' warehouses.¹⁹⁶ Losses in film and music are equally dramatic. Only about twenty percent of feature films from the 1920s and fifty percent of feature films made prior to 1950 still survive, and it is estimated that sixty percent of all sound recordings are now not commercially available.¹⁹⁷

While the first sale doctrine cannot be relied on for printing or reprinting works that are out of print, it plays an important role in mitigating the potential cultural loss associated with works that go out of print. Exhaustion rules open up the possibility of a secondary

194. Hovenkamp, *supra* note 20, at 541.

195. Deirdre K. Mulligan & Jason M. Schultz, *Neglecting the National Memory: How Copyright Term Extensions Compromise the Development of Digital Archives*, 4 J. APP. PRAC. & PROCESS 451, 459 (2002).

196. Diane Leenheer Zimmerman, *Can Our Culture Be Saved? The Future of Digital Archiving*, 91 MINN. L. REV. 989, 1003 (2006).

197. Eric Matthew Hinkes, *Access Controls in the Digital Era and the Fair Use/First Sale Doctrines*, 23 SANTA CLARA COMPUTER & HIGH TECH. L.J. 685, 702 (2006).

market and assure that the artifacts embedding protected works remain available to the public over time.¹⁹⁸

In addition, exhaustion rules help ensure cultural preservation by reducing the opportunity cost and risk of acquiring an intellectual work in the first place. Consumers may be more willing to purchase an item if they know they can later resell it in a secondary market¹⁹⁹ (e.g., to a used or antiquarian bookseller) without being required to obtain the copyright owner's permission.

Finally, the first sale doctrine contributes directly to the survival of copies of works or patented goods over time by discouraging abandonment and waste: instead of discarding an item when keeping or preserving it is costly and the item is no longer useful, convenient, or economic to keep, the first sale doctrine makes it legal to sell or donate a used copy of an intellectual good.²⁰⁰ In short, the first sale doctrine enshrines preference for the garage sale over the garbage bin and for the library over the landfill.

In addition to these static benefits, the first sale doctrine contributes to dynamic efficiency by permitting secondary market channels that enable works and the ideas they carry, or goods and the technologies embedded therein, to remain accessible to the public even if the copyright holder ceases production or distribution of the work.²⁰¹ It creates a built-in back-up system that ensures public access to history, culture, and technology, even if the IP holder and the immediate buyer or licensee of her work rationally ignore, or strategically choose to negate, this benefit; it preserves the public bargain that underlies IP law, even if the private bargain between the IP owner and its transacting party ignores it.

198. R. Anthony Reese, *The First Sale Doctrine in the Era of Digital Networks*, 44 B.C. L. REV. 577, 592 (2003); see generally Diane L. Zimmerman, *Cultural Preservation: Fear of Drowning in a Licensing Swamp*, in WORKING WITHIN THE BOUNDARIES OF INTELLECTUAL PROPERTY (Rochelle C. Dreyfuss, Diane L. Zimmerman & Harry First eds., 2010); Hinkes, *supra* note 197, at 685; Margaret Jane Radin, *Regulation by Contract, Regulation by Machine*, 160 J. INST. & THEORETICAL ECON. 1 (2004); Zimmerman, *supra* note 196, at 989.

199. Reese, *supra* note 198, at 607.

200. *Id.* at 607–08.

201. Hinkes, *supra* note 197, at 689.

B. Protecting the “Innovation Wetlands”

Implicit in the debate about exhaustion is the argument that whatever benefits IP owners ultimately benefits society because these benefits provide greater incentives to create and greater benefits to disseminate intellectual goods. An extreme version of this argument would be that even if post-sale restraints are imposed for the sole reason of increasing the IP owner’s market power, there is nothing objectionable about that because the additional profit will benefit society in the long run.²⁰² Such a view is consistent with the predominantly producer-centric model of innovation. In this model, the most important designs or innovations originate from producers and are supplied to consumers through the sale of goods and services.²⁰³ The higher the profit available, the greater is the incentive to innovate.²⁰⁴ However, recent research on user innovation challenges this nostrum.²⁰⁵ It shows that innovation often occurs outside the producer-firm, suggesting that sensible innovation policies may require more than just ex ante justification; they must also focus on conditions that emerge ex post and protecting the “Innovation Wetlands,” a term coined by Andrew Torrance and Eric Von Hippel.²⁰⁶

The producer-centric model of innovation, which dominates current debates around IP and antitrust, confronts a moving empirical reality. As Eric von Hippel observes, “innovation by individual users . . . and . . . open collaborative innovation, are

202. *Cf. Verizon Commc’ns, Inc. v. Law Offices of Curtis V. Trinko, LLP*, 540 U.S. 398, 407 (2004) (“The mere possession of monopoly power, and the concomitant charging of monopoly prices, is not only not unlawful; it is an important element of the free-market system. The opportunity to charge monopoly prices—at least for a short period—is what attracts ‘business acumen’ in the first place; it induces risk taking that produces innovation and economic growth.”).

203. Carliss Baldwin & Eric von Hippel, *Modeling a Paradigm Shift: From Producer Innovation to User and Open Collaborative Innovation*, 22 *ORG. SCI.* 1399 (2011).

204. *Id.*

205. *Id.*

206. Andrew W. Torrance & Eric von Hippel, *Protecting the Right to Innovate: Our ‘Innovation Wetlands’* 3–4 (Oct. 9, 2013), *available at* <http://ssrn.com/abstract=2339132> (“We define the innovation wetlands as the rights and conditions that enable innovation by individuals to flourish. Just as in the case of environmental wetlands, the nature and extent of the innovation wetlands must be understood, and the value of the considerable innovation activity that takes place therein must be better appreciated.”).

modes of innovating that increasingly compete with and may displace producer innovation in many parts of the economy.”²⁰⁷ Thus, there is not a single model of innovation. There are three: one driven by producers, one facilitated by users, and one created through open collaboration.

User innovation is driven by lead users having an intention to use rather than sell an innovative technology. It arises where design and transaction costs are low. The distinguishing characteristic of this model is that lead users typically reveal their designs for free, without profit motive, at least not an immediate one.²⁰⁸ This makes user innovation more desirable over alternative methods of innovation when it is technologically feasible; its openness imposes less deadweight loss on consumers, while minimizing transaction costs for other future producers.²⁰⁹

Many user innovators are lead users. These users have heterogeneous needs not easily satisfied by mass-market production.²¹⁰ Their unique intellectual capital and idiosyncratic needs and experiences often allow them to encounter problems and devise solutions that no one else can.²¹¹ They frequently anticipate features that improve products but for which no general demand yet exists. Lead users are thus distinguishable from the general consumer in that they possess both a need for optimization that impels modification of an existing product for their own use and an expertise to accomplish this modification.²¹² The common feature among user innovations is that they emerge from many small ideas,

207. Baldwin & von Hippel, *supra* note 203.

208. Dietmar Harhoff, Joachim Henkel & Eric von Hippel, *Profiting from Voluntary Information Spillovers: How Users Benefit by Freely Revealing Their Innovations*, 32 RES. POL'Y 1753 (2003).

209. Baldwin & von Hippel, *supra* note 203, at 1414.

210. ERIC VON HIPPEL, *DEMOCRATIZING INNOVATION* 34–43 (2005).

211. Leah M. Theriault, *A User Innovation Theory of the Numerus Clausus* 51–66 (2011) (unpublished S.J.D. thesis, University of Toronto), *available at* https://tspace.library.utoronto.ca/bitstream/1807/35743/3/Theriault_Leah_M_2011June_SJD_thesis.pdf; *see also* SCOTT E. PAGE, *THE DIFFERENCE* 31, 153–54 (2007) (discussing the relevance of the individual's perspective, heuristics, and a toolbox to solving innovative problems); VON HIPPEL, *supra* note 210, at 147–64 (discussing users having different “toolkits,” which allow them to match their skills to a particular problem).

212. Theriault, *supra* note 211, at 57–58.

obtained through use and manipulation of an asset, over long periods of time.²¹³

If non-producer-centric models of innovation are becoming important sources of innovation, then the design of optimal exhaustion rules should take into account these other sources of innovation. In the case of user innovation, exhaustion fosters such innovation because it allows a user who possesses a copy or an artifact to experiment with it without the need to obtain permission from the IP owners, and it further fosters user innovation because it permits users who do not innovate to resell or otherwise transfer the goods to users who do innovate.²¹⁴

Implied in the view that IP owners should be free to opt out of the first sale doctrine is the notion that inventors or authors must be able to capture the full social benefit of their innovation for there to be adequate incentive to innovate,²¹⁵ and that unless and until an innovator fully internalizes the cost of innovation, he or she will not be able to efficiently manage an innovation after it is created.²¹⁶ The implied presumption is that the inventor is the most informed about the value of an invention, and that value can be captured fully by the transaction price.

User innovation challenges this thesis because it shows that buyers and sellers may have very little idea at the time of invention or at the point of sale what innovations will ultimately emerge. If innovation is tied to asset use and emerges therefrom, it is impossible for the seller to account fully for the asset's value at the time of sale. Few can predict what innovation will follow. Fewer still can predict what value such innovation may have for social, cultural, and political life.²¹⁷ The best the seller can do is to price the asset according to a

213. R. Keith Sawyer, *Creativity, Innovation, and Obviousness*, 12 LEWIS & CLARK L. REV. 461, 479 (2008).

214. See LAWRENCE LESSIG, *THE FUTURE OF IDEAS: THE FATE OF THE COMMONS IN A CONNECTED WORLD* 12 (2001) (discussing innovation without permission).

215. See, e.g., Hovenkamp, *supra* note 20, at 524, 529.

216. This view is consistent with the work of Harold Demsetz on externalities. See Harold Demsetz, *Toward a Theory of Property Rights*, 57 AM. ECON. REV. 347 (1967); WILLIAM J. BAUMOL, *THE FREE MARKET INNOVATION MACHINE* 5 (2002). For a critique of this argument, see Frischmann & Lemley, *supra* note 188, at 268–71.

217. YOCHAI BENKLER, *THE WEALTH OF NETWORKS: HOW SOCIAL PRODUCTION TRANSFORMS MARKETS AND FREEDOMS* 129–355 (2006) (discussing how open collaborative

cost-benefit analysis at a static point in time, a time where the probabilities of benefits are innately unknowable and will be seriously discounted by the seller and the buyer.

User innovation also implies that knowing who is best able to maximize the value of an asset at the time of transaction is unlikely. This suggests that the best-informed party, the most capable steward of innovation, may be neither the IP owner who imposes the restraint nor the firm with which she transacts and accepts the restraint. Moreover, the innovator may not even exist when the terms of the transaction between the IP owner and its transacting parties are negotiated. Therefore, in a truly Coasian way, if we cannot tell *ex ante* who is in the best position to further innovate, rules allowing possessors of goods to innovate without restraints reduce transaction costs.²¹⁸

Restrictions on post-sale use are often attempts by producers to internalize or capture as much of the social benefit from an innovation as possible through contract or other means.²¹⁹ But attempts to internalize such social benefits necessarily imply limiting the flow of spillovers. Contrary to the view that optimal innovation requires full-appropriation, there is evidence that the spillovers from innovation actually encourage additional innovation.²²⁰ Accordingly, the most efficient means of fostering innovation may be to allow knowledge spillovers by inhibiting post-sale restraints rather than allowing them. As Lemley and Frischmann explain, “[i]f a technology might be repurposed or improved in ways that the initial

innovation facilitates new social and political realities as well).

218. Support for this conclusion may be found in the network neutrality literature, for example, where innovation is said to proliferate precisely because the Internet is a general purpose technology whose end-to-end architecture eliminates transaction costs among user innovators. See RICHARD G. LIPSEY, KENNETH I. CARLAW & CLIFFORD BEKAR, *ECONOMIC TRANSFORMATIONS: GENERAL PURPOSE TECHNOLOGIES AND LONG-TERM ECONOMIC GROWTH* 1–219 (2005) (describing growth, technological change, and general purpose technologies); BARBARA VAN SCHEWICK, *INTERNET ARCHITECTURE AND INNOVATION* 67–111, 357–61 (2010) (analyzing the end-to-end argument’s relationship to innovation); JONATHAN ZITTRAIN, *THE FUTURE OF THE INTERNET AND HOW TO STOP IT* 70 (2008) (discussing the importance of “generative” systems for innovation).

219. Hovenkamp, *supra* note 20, at 524.

220. Frischmann & Lemley, *supra* note 188, at 268.

innovator is unlikely to foresee, full internalization will interfere with the socially optimal development and use of a technology.”²²¹

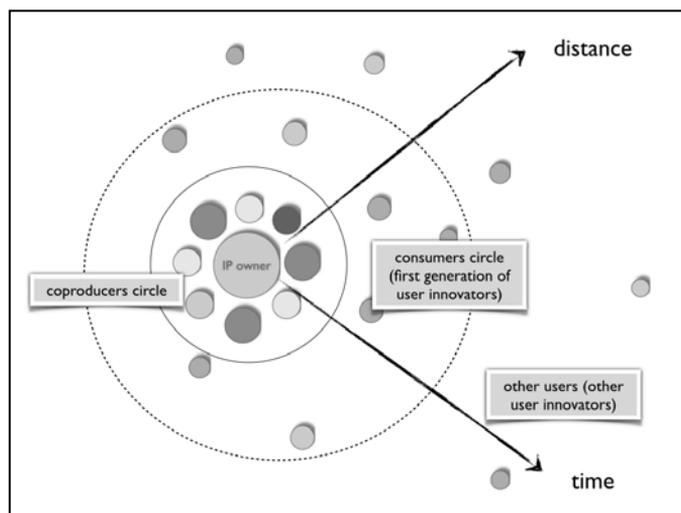


Figure 2 An IP owner, coproducers, and user innovators

Figure 2 graphically depicts this assertion. Recognizing that users are not only consumers but also actual or potential innovators, implies that granting IP owners an extended power (over time and distance) to restrain the use of goods embodying their innovation will impede users’ ability to innovate, or transfer the goods to others who might innovate. A sticky first sale doctrine prevents such impedance and preserves freer grounds for future innovation.

In sum, the first sale doctrine improves welfare by facilitating the efficient use of physical and intellectual assets (including human capital) both statically and dynamically. However, IP owners and their immediate transacting parties may rationally impose and agree to restraints that maximize their short-term private benefits, while ignoring the short—and long—term externalities that such restraints may generate. Full notice of the restraint does not remedy the problem, because notice only guarantees more efficient bargaining between the transacting parties, but does not account for the

221. *Id.* at 278.

externalities. If the terms of the transaction between the IP owner and her transacting party are unlikely to reflect what is socially optimal, contracting around the first sale doctrine should be met with a healthy dose of legal skepticism. Therefore, the burden of demonstrating the efficiency of the restraints should fall on the transacting parties. When the restraints purport to bind third parties or have long-term effects, the dose of suspicion should increase.

IX. APPLICATION

This Part demonstrates how the framework developed in this Article bears on recent cases involving the first sale doctrine.

A. Costco v. Omega

Costco v. Omega was a parallel trade case that conforms to the paradigmatic international price-discrimination story.²²² Costco purchased Omega watches, originally sold by Omega to distributors in Egypt and Paraguay, from a third-party importer.²²³ Costco sold the watches for \$1,300 each instead of the \$2,000 suggested retail price. Omega sued Costco for copyright infringement.²²⁴ The copyright in question was in a logo engraved on the back of the watch.²²⁵ The Ninth Circuit agreed with Omega, holding that the first sale doctrine did not apply because under 17 U.S.C. § 109(a) the right to resell that copy without the authority of the copyright owner is limited to copies “lawfully made under this title,” which the court interpreted as lawfully made in the United States.²²⁶ Because the copies were made outside the United States, the United States distribution right had not been exhausted by the authorized foreign sales, and the unauthorized importation constituted infringement of the copyright.²²⁷

222. *Omega S.A. v. Costco Wholesale Corp.*, 541 F.3d 982 (9th Cir. 2008), *aff’d by an equally divided court*, 131 S. Ct. 565 (2010).

223. *Id.* at 984.

224. *Id.*

225. *Id.* at 985.

226. *Id.* at 989 (quoting *Quality King Distribs., Inc. v. L’anza Research Int’l, Inc.*, 523 U.S. 135, 148 (1998)).

227. *Id.*

An equally divided Supreme Court affirmed the Ninth Circuit judgment,²²⁸ but ultimately Costco prevailed.²²⁹ On remand, the district court held that Omega's attempt to prevent parallel importation of noncopyrightable watches, by suing the importer for unlawfully importing the copyrighted company logo embedded on the back of the watch, constituted copyright misuse.²³⁰ Moreover, the court also decided to award Costco reasonable attorney fees, because, among other reasons, Omega's infringement action "was arguably unreasonable and frivolous . . . [and] clearly not one properly raised under copyright law."²³¹ Copyright law "serves the purpose of enriching the general public through access to creative works," the court reasoned, but "[b]y affixing a barely perceptible copyrighted design to the back of some of its watches, Omega did not provide—and did not seek to provide—creative works to the general public. Omega sought to exert control over its watches, control which it believed it could not otherwise exert."²³²

To some extent, *Omega* was an easy case for proponents of international exhaustion and national exhaustion alike. For proponents of international exhaustion, a case like *Omega* illustrates the dangers of national exhaustion, and how a copyright ban on parallel imports can be misused to restrain trade in noncopyrightable products, such as watches or chocolate bars.²³³ For proponents of national exhaustion, Costco's happy ending can demonstrate how doctrines such as copyright misuse, fair use, or implied license, can be relied on to prevent those misuses. In that regard, *Omega* was somewhat of a red herring, deflecting attention away from the real dilemma that underlies the tension between international and national exhaustion. The Supreme Court would have another opportunity to resolve this dilemma in 2013, in *Kirtsaeng v. John Wiley & Sons*.²³⁴

228. Costco Wholesale Corp. v. Omega, S.A., 131 S. Ct. 565 (2010).

229. See Omega S.A. v. Costco Wholesale Corp., No. CV 04-05443 TJH, 2011 WL 8492716, at *2 (C.D. Cal. Nov. 9, 2011).

230. *Id.*

231. Omega S.A. v. Costco Wholesale Corp., No. CV 04-05443 TJH (RCx), 2012 WL 3150432, at *1 (C.D. Cal. June 20, 2012).

232. *Id.* (quoting Fogerty v. Fantasy, Inc., 510 U.S. 517, 527 (1994)).

233. Euro-Excellence Inc. v. Kraft Canada Inc., [2007] 3 S.C.R. 20 (Can.).

234. 133 S. Ct. 1351 (2013).

B. Kirtsaeng v. John Wiley & Sons, Inc.

Supap Kirtsaeng, a citizen of Thailand, moved to the United States in 1997 to study mathematics at Cornell University.²³⁵ His education was paid for with the help of a Thai government scholarship which required him to teach in Thailand for 10 years on his return. Kirtsaeng successfully completed his undergraduate courses at Cornell and a Ph.D. program in mathematics at the University of Southern California, and then returned to Thailand to teach. While he was studying in the United States, Kirtsaeng asked his friends and family in Thailand to buy copies of foreign edition English-language textbooks, which were sold at lower prices at Thai book shops, and mail them to him in the United States. He then sold them, reimbursed his family and friends, and kept the profit, repeating this process over and over.²³⁶

John Wiley & Sons, the publisher of some of these books, brought an action for copyright infringement against Kirtsaeng, claiming that Kirtsaeng's unauthorized importation of its books and his later resale of those books infringed Wiley's § 106(3) exclusive right to distribute as well as § 602(a)'s related import prohibition.²³⁷ Kirtsaeng replied that the copies that he imported and sold were lawfully made and therefore protected under the first sale doctrine. The question, therefore, was whether the first sale doctrine, as codified in § 109(a), applies to copies that were made abroad. It was not disputed that the copies were lawfully made, but like the Ninth Circuit in *Omega*, the lower courts in this case held that "lawfully made under this title" meant "lawfully made in the United States."²³⁸

The majority held that § 109(a) does not restrict the scope of the first sale doctrine to copies made in the United States, and that the doctrine equally applies to copies of a copyrighted work lawfully made abroad.²³⁹ Justice Ginsburg, joined by Justices Kennedy and

235. *Id.* at 1356.

236. *Id.*

237. *Id.* at 1357.

238. *Id.* at 1358 (citing *Denbicare U.S.A. Inc., v. Toys "R" Us, Inc.*, 84 F.3d 1143, 1149–50 (9th Cir. 1996), *abrogated by Kirtsaeng*, 133 S. Ct. 1351).

239. *Id.*

Scalia, issued a long and passionate dissent. “Instead of adhering to the Legislature’s design,” she accused the Court of adopting “an interpretation of the Copyright Act at odds with Congress’s aim to protect copyright owners against the unauthorized importation of low-priced, foreign-made copies of their copyrighted works.”²⁴⁰ She characterized the Court’s opinion as a “bold departure from Congress’ design [which] is all the more stunning, for it places the United States at the vanguard of the movement for ‘international exhaustion’ of copyrights—a movement the United States has steadfastly resisted on the world stage.”²⁴¹

Justice Kagan, joined by Justice Alito, concurred with the majority opinion, but intimated that while the first sale doctrine is not limited geographically, in enacting § 602(a)(1) Congress probably intended “to permit market segmentation, . . . not by removing first sale protection from every copy manufactured abroad . . . but by enabling the copyright holder to control imports even when the first sale doctrine applies”²⁴² She recognized, however, that the Court’s earlier holding in *Quality King*²⁴³ had rendered this interpretation impossible.²⁴⁴

Whether the Court correctly interpreted § 109(a) and § 602(a)(1) exceeds the scope of this Article. Admittedly, each of the interpretations requires some linguistic acrobatics, which probably reflects the contentious nature of the issue of parallel imports. I would like, however, to offer a few observations. The first relates to the immediate issue of parallel imports, and the other comments relate to the first sale doctrine more generally.

The immediate issue in *Kirtsaeng* was the application of the first sale doctrine to copies that were made abroad, and since the case involved parallel imports, it was presented as demanding a choice between national and international exhaustion, as well as about the fate of market segmentation and cross-country price discrimination. Notably, however, despite the dissent’s strong conviction that Congress intended to implement national exhaustion in order to

240. *Id.* at 1373 (Ginsburg, J., dissenting).

241. *Id.* (Ginsburg, J., dissenting).

242. *Id.* (Kagan, J., concurring).

243. *Quality King Distribs., Inc. v. L’anza Research Int’l, Inc.*, 523 U.S. 135 (1998).

244. 133 S. Ct. at 1372 (Kagan, J., concurring).

promote market segmentation and price discrimination, the statutory hook for that conclusion—the argument that the phrase “made under this title” renders the first sale doctrine inapplicable to copies made outside the United States—is rather peculiar: what matters for market segmentation is not where the copy was made, but where the copy was initially distributed.²⁴⁵ Presumably, if Congress were unequivocally interested in assisting copyright owners to segment markets internationally by legislating a national exhaustion rule, it would have done that on the basis of where the copy was sold, not on the basis of where it had been made.

Indeed, a legal rule that treats the circulation of goods differently depending on the question of where the copy of a work was made may create perverse incentives. Instead of deciding where to produce copies on the basis of economic factors, firms may shift production of copies abroad in order to gain greater control over the downstream distribution of their goods. Whether such control is desirable or not, it should be granted or denied regardless of the place where the copy was made.²⁴⁶ If Congress clearly intended to assist copyright owners to segment international markets by preventing parallel imports, it could have done so in a more straightforward and elegant way; for example, by adding the word “import” in § 106(3). Doing that would have created a clear right of importation, which could then be exhausted pursuant to § 109(a), all the while being neutral to the location of the copy’s origin (and likewise, if Congress clearly intended to adhere to international exhaustion, it could have done that easily).²⁴⁷

Second, the dissent’s proffered rationale for favoring national exhaustion is the supposed benefits of price discrimination. However, as discussed in Part V.C above, even if price discrimination is desirable, there is little economic justification to treat domestic price discrimination differently from international price

245. *Id.* at 1373 n.2.

246. *Id.*

247. For example, § 109(a) could read: Notwithstanding the provisions of section 106(3), the owner of a particular copy or phonorecord *that was made with the authority of the copyright owner wherever it was made*, or any person authorized by such owner, is entitled, without the authority of the copyright owner, to sell or otherwise dispose of the possession of that copy or phonorecord.

discrimination. If IP law should be harnessed to enforce price discrimination schemes and prevent arbitrage, it makes little sense to target only cross-border arbitrage, but deny remedy from all other forms of domestic arbitrage. Therefore, the only intellectually coherent choice is not between national and international exhaustion, but rather between universal exhaustion and no exhaustion at all.²⁴⁸

Third, the effect of international exhaustion on the extent of international market segmentation may be marginal at best. On the one hand, in some cases publishers may refrain from price discriminating even if the law allows them to prevent arbitrage;²⁴⁹ and on the other hand, publishers may resort to implement price discrimination even if they have no legal recourse against any arbitrageur.²⁵⁰ Moreover, parallel imports themselves are a form of price discrimination between consumers within the importing market, namely those who are willing to pay the ‘official’ high price, and those who are not and thus seek a bargain.²⁵¹

Furthermore, it is doubtful that a parallel importation enterprise such as Kirtsaeng’s can have such a significant impact on sales in the United States to prompt any significant change in the publisher’s international distribution system. Recall that Mr. Kirtsaeng’s supply chain consisted of family and friends who would purchase the books in Thailand and ship them to the United States. An operation like this can probably be carried on efficiently only on a limited scale, beyond which the marginal cost of obtaining additional copies and shipping them to the United States is likely to increase very quickly. The more general point is that parallel importers operate at a cost disadvantage compared to the official distribution system, which caps the growth of their operations.

These observations are consistent with the Court’s holding that the fact a publisher may find it more difficult to charge different

248. See *supra* Part V.C.

249. See Katz, *Network Effects*, *supra* note 71, at 181 (explaining that publishers may refrain from price discrimination even if arbitrage could be blocked, because information about legal copies that are sold elsewhere for a significantly lower price cannot be blocked, and such information is a signal that might affect consumers’ preferences and their willingness to pay the current higher price).

250. Rub, *supra* note 88, at 45.

251. Reza Ahmadi & B. Rachel Yang, *Parallel Imports: Challenges from Unauthorized Distribution Channels*, 19 *MARKETING SCI.* 279, 279 (2000).

prices for the same book in different geographic markets is beside the point because there is “no basic principle of copyright law that suggests that publishers are especially entitled to such rights.”²⁵² To the contrary, the first sale doctrine is a conscious limitation on the ability to price discriminate.²⁵³

Kirtsaeng helps illuminate a few of the other points developed earlier in this article. It is noteworthy that the entire Court seems to have been troubled by the “parade of horrors”—a long list of allegedly absurd outcomes that could result from the plaintiff’s theory that the first sale doctrine was limited to goods manufactured in the United States.²⁵⁴ Indeed, the “parade of horrors” is a manifestation of the problem of post-sale restraints that can be enforced as property rules, and that continue to encumber products in which IP is embedded even when those restraints outlive their useful purpose.²⁵⁵

As Parts VI to VIII demonstrate, while allowing IP owners to invoke IP rights to tightly control the distribution of their goods may improve the efficiency of their distribution systems, IP remedies are not generally essential for achieving such efficiency because contractual remedies might be sufficient. It is not enough to show that relying on IP remedies is more efficient than relying on contractual remedies. To justify the availability of IP remedies it must be shown that the marginal benefit of IP remedies over contractual remedies outweighs the social cost that is associated with IP but lacking with contract. More specifically, there is no basis to assume that maintaining an efficient distribution system for textbooks requires preventing their free circulation for decades, until the copyright therein expires.

The majority solved the problem by opting for international exhaustion, emphasizing the common law (rather than statutory) origin of the doctrine, and evincing strong commitment to the continuing relevance of the common law’s hostility to restraints on alienation.²⁵⁶

252. *Kirtsaeng v. John Wiley & Sons, Inc.*, 133 S. Ct. 1351, 1370. (2013).

253. *Id.* at 1371.

254. Rub, *supra* note 88, at 47.

255. *See supra* Part VI.A.

256. *Kirtsaeng*, 133 S. Ct. at 1363.

The concurring opinion suggested a partial solution, which would give copyright owners the power to prevent unauthorized importation, but without imposing downstream liability on those who “purchase and resell in the United States copies that happen to have been manufactured abroad.”²⁵⁷ This solution prevents the absurd outcome that the distribution right *never* exhausts with respect to a copy made abroad, and “would target unauthorized importers alone, and not the ‘libraries, used-book dealers, technology companies, consumer-goods retailers, and museums’ with whom the Court today is rightly concerned.”²⁵⁸ However, while Justice Kagan’s solution might prevent some of the specific horrors with which the Court was rightly concerned (e.g., the specter that libraries would need the copyright owner’s permission for *lending* a foreign made book, or that museums would need to obtain clearance before they could *display* a work of art made abroad), it does not eliminate all horrors, because the *importation* of any work,²⁵⁹ or any other product embodying a copy of a work, would still require the copyright owner’s permission. Thus, the concerns of used-book dealers, technology companies, consumer-goods retailers, and museums remain intact. Justice Kagan’s solution might prevent some of the horrors attaching to the downstream transfer or use of imported copies *after* they have been circulated in the United States—regardless of whether their importation had been authorized or not²⁶⁰—but it also means that for the entire duration of the

257. *Id.* at 1372–73.

258. *Id.* at 1373.

259. Or in the case of libraries, the importation of more than the five copies explicitly permitted under 17 U.S.C. § 602(a)(3)(C).

260. It is not entirely clear whether Justice Kagan’s solution envisions protection only for those who buy and sell copies whose importation was authorized by the copyright owner, or anyone who buys or sells such copies in the United States, irrespective of whether their importation was lawful or not. If she envisions the latter option, it is not entirely clear that such protection would be afforded to them. Selling such copies may be protected under § 109, but since under § 501 the importation of such copies is an act of infringement, such copies are infringing copies, and it is not entirely clear that the copyright owner does not have any recourse against those who possess such infringing copies. In *Societe Civile Succession Richard Guino v. Int’l Found. for Anticancer Drug Discovery*, 460 F. Supp. 2d 1105 (D. Ariz. 2006) the court ultimately held that § 503(a) does not authorize the impoundment of infringing property purchased by a non-infringing person, but later decisions recognized that this is not necessarily a foregone conclusion. *See Virtual Studios, Inc. v. Beaulieu Grp., LLC*, — F.Supp.2d —, 2013 WL 6629040, at *10 (E.D.Tenn. 2013).

copyright, copyright owners have complete power to prevent those copies from entering the United States in the first place. This means that the cultural, scientific, and technological choices of Americans would be strictly limited to purchasing only those works that are commercially available in the United States. Importing any other work (including second-hand books, out-of-print books, books in foreign language or on niche topics that may be only available abroad), or any product embodying such work, would require the copyright owner's permission, even when lawfully made and lawfully sold abroad.

Justice Kagan seems to be sympathetic to the notion that § 602(a)(1) was indeed designed to give copyright holders greater power to segment international markets by restricting importation. The problem, however, is that her proposed solution—prohibiting any unauthorized importation of lawfully made copies—might prevent from entering the United States any copy of a lawfully made work even when there is no distribution system that the importation might undermine. This would be another example of an IP remedy that may be available even when it has outlived its useful purpose.

The dissent's solution is interesting, too. While it is adamant that Congress unequivocally intended to implement national exhaustion, it ignores that if that were the case, Congress did so in a rather awkward way—by basing exhaustion on the place of manufacturing rather than on the country of first sale. But it also acknowledges some of the absurdities arising from this interpretation. To mitigate them, the dissent offers two moves: one is the suggestion that possible defenses, such as fair use or implied license, may mitigate some of the “horribles”;²⁶¹ the other is the suggestion that the logic of *Bobbs-Merrill* and the common law first sale doctrine can be used to apply the first sale doctrine to foreign made copies *after* their first authorized importation into the United States.²⁶²

As I noted above, while doctrines such as fair use, implied license, or misuse, can solve some of the problems associated with unexhausted restraints, they do not fully solve the problem because

261. *Kirtsaeng*, at 1389.

262. *Id.* at 1386–87.

they shift the burden of invalidating the restraint on the user.²⁶³ I believe, however, that the fact that both the majority and the dissent confirmed the existence of a common law first sale doctrine whose application is not necessarily bound by the text of § 109(a) may prove to be an important development that might guarantee the doctrine's continued viability. The anchoring of the first sale doctrine in the common law rather than the in the strictures of § 109(a) might prove important in the case copyrighted works in digital form, as it might prevent some problematic refusals of courts to apply the doctrine in the context of digital works, such as in the following case, *Vernor v. Autodesk*,²⁶⁴ and *Capitol Records v. ReDigi*,²⁶⁵ discussed later.

C. Vernor v. Autodesk

Vernor v. Autodesk involved the resale of used copies of software. Timothy Vernor purchased several used copies of Autodesk, Inc.'s AutoCAD Release 14 software ("Release 14") from one of Autodesk's direct customers and then resold the copies on eBay. Vernor then brought a declaratory judgment action against Autodesk to establish that the resale did not infringe Autodesk's copyright.²⁶⁶ Vernor argued that the first sale doctrine permits the sale of those used copies.²⁶⁷ The Ninth Circuit disagreed.²⁶⁸ The court reasoned that "[i]n its current form, [the first sale doctrine] allows the 'owner of a particular copy' of a copyrighted work to sell or dispose of his copy without the copyright owner's authorization."²⁶⁹ The question, therefore, was whether Autodesk's direct customer (who sold the

263. *See supra* Part VI.A.

264. *Vernor v. Autodesk, Inc.*, 621 F.3d 1102, 1111 (9th Cir. 2010).

265. *Capitol Records, LLC v. ReDigi Inc.*, 934 F. Supp. 2d 640 (S.D.N.Y. 2013).

266. *Vernor*, 621 F.3d at 1111. On several earlier occasions, Vernor sought a declaratory judgment after Autodesk filed Digital Millennium Copyright Act ("DMCA") take-down notices with eBay claiming that Vernor's sale infringed its copyright and eBay terminated Vernor's auctions. *Id.* at 1105–06. In each of these cases the auctions were reinstated after Vernor filed counter-notices and Autodesk failed to respond. *Id.* At one point, however, eBay suspended Vernor's account for a month because of Autodesk's repeated allegations of infringement, during which Vernor was unable to earn any income from eBay. *Id.* at 1106.

267. *Id.* at 1106.

268. *Id.*

269. *Id.* at 1107 (quoting 17 U.S.C. § 109(a) (2006)).

copy to Vernor) was an “owner of a particular copy.”²⁷⁰ The court found that the customer was not an owner of the copy, but instead a licensee because the Software License Agreement that the customer had to accept before installing the software: “(1) specific[d] that the user is granted a license,” while ownership in the copy remained with Autodesk; “(2) significantly restrict[ed] the user’s ability to transfer the software; and (3) impose[d] notable use restrictions.”²⁷¹ Finally, the court noted “the significant policy considerations raised by the parties and amici on both sides of this appeal,”²⁷² but decided that as a matter of statutory interpretation and precedent, such

270. *Id.*

271. *Id.* at 1111.

272. *Id.* at 1114. For example, Autodesk and its supporters argued that “judicial enforcement of software license agreements that restrict transfers of copies of the work” is desirable because it “(1) allows for tiered pricing for different software markets, such as reduced pricing for students or educational institutions; (2) increases software companies’ sales; (3) lowers prices for all consumers by spreading costs among a large number of purchasers; and (4) reduces the incidence of piracy by allowing copyright owners to bring infringement actions against unauthorized resellers.” *Id.* at 1114–15. On the other hand, Vernor, eBay, and the American Library Association (“ALA”) have presented policy arguments against the court’s decision. *Id.* at 1115. Vernor argued that the decision “(1) does not vindicate the law’s aversion to restraints on alienation of personal property; (2) may force everyone purchasing copyrighted property to trace the chain of title to ensure that a first sale occurred; and (3) ignores the economic realities of the relevant transactions, in which the copyright owner permanently released software copies into the stream of commerce without expectation of return in exchange for upfront payment of the full software price.” *Id.* Additionally, eBay argued in favor of “a broad view of the first sale doctrine,” which necessarily “facilitate[s] the creation of secondary markets for copyrighted works” and “contributes to the public good by (1) giving consumers additional opportunities to purchase and sell copyrighted works, often at below-retail prices; (2) allowing consumers to obtain copies of works after a copyright owner has ceased distribution; and (3) allowing the proliferation of businesses.” *Id.* Moreover, the ALA augmented eBay’s argument by stating “that the first sale doctrine facilitates the availability of copyrighted works after their commercial lifespan, by *inter alia* enabling the existence of libraries, used bookstores, and hand-to-hand exchanges of copyrighted materials.” *Id.* It also argued “that judicial enforcement of software license agreements, which are often contracts of adhesion, could eliminate the software resale market, require used computer sellers to delete legitimate software prior to sale, and increase prices for consumers by reducing price competition for software vendors.” *Id.* In response to Autodesk’s arguments, the ALA asserted that were it to be upheld it would “(1) undermine[] 17 U.S.C. § 109(b)(2), which permits non-profit libraries to lend software for non-commercial purposes, and (2) would hamper efforts by non-profits to collect and preserve out-of-print software.” *Id.* The ALA’s position reflects its “fears that the software industry’s licensing practices could be adopted by other copyright owners, including book publishers, record labels, and movie studios.” *Id.*

considerations should be left to Congress.²⁷³ As a result, by requiring users to agree to the Software License Agreement, Autodesk was able to impose nonexhausting post-sale restraints on copies of its software that bind all future users.

To determine whether this outcome is desirable, it is useful to understand what Autodesk tried to achieve by preventing the resale of its software. The most likely reason to prevent resale is to create a price discrimination scheme. For example, according to the court, Autodesk offers the software with different terms for commercial, educational institution, and student users.²⁷⁴ Arguably, maintaining this scheme requires an ability to prevent arbitrage (e.g., preventing a student user from reselling her copy to a commercial user), and preventing arbitrage is easier when the unauthorized sale of a lawfully-made copy constitutes copyright infringement.

The case, however, did not involve a student or an educational user reselling her copy to a commercial user, but rather dealt with a commercial user upgrading his software to a newer version and reselling the old one. This raises the possibility that Autodesk sought to prevent the resale to maintain another type of price discrimination. When Autodesk released AutoCAD Release 15 it sold it for \$3,750.²⁷⁵ A commercial user new to AutoCAD would have to pay this amount if Release 15 was the only suitable software for her needs.²⁷⁶ However, an existing user of Release 14 could continue using the earlier version even if the new version was better. In other words, existing users have more elastic demand for the new version than new users. In order to convince existing users to upgrade, Autodesk offered them the new version for a substantially lower price: \$495.²⁷⁷ If the upgrading users could then resell the older version to others, Autodesk's new version would no longer be the only game in town. Rather than paying \$3,750, a new user

273. *Id.* at 1115.

274. *Id.* at 1104.

275. *Id.* at 1105.

276. I assume away the availability of competing programs. This assumption is not unrealistic, as AutoCAD is currently the de facto standard in this type of programs, with a market share of eighty-five percent. See *Autodesk (ATSK)*, WIKINVEST, [http://www.wikininvest.com/stock/Autodesk_\(ADSK\)](http://www.wikininvest.com/stock/Autodesk_(ADSK)) (last visited Feb. 14, 2014).

277. *Vernor*, 621 F.3d at 1105.

would have the option of buying a used copy of the older version at a low price and then upgrading to Release 15 for Autodesk's existing user price of \$495.

The welfare implications of such a scheme are ambiguous. Vernor's engagement in the resale of used copies indicates that the quantity sold and prices set by Autodesk do not satisfy the entire demand for AutoCAD (otherwise, there would be no demand for the copies sold by Vernor), but it is far from clear that by eliminating the secondary market Autodesk will increase output and satisfy this demand. Therefore, it is not clear whether allowing Autodesk to control the secondary market increases or decreases allocative efficiency. There is no clear answer about the effect on dynamic efficiency. Evidently, Autodesk's attempts to eliminate Vernor strongly suggest that the secondary market eats into its profit, and this can be a bad thing if lower profits reduce the incentive to innovate. On the other hand, it is equally possible that competition from the secondary market drives Autodesk to keep innovating, rather than selling the same version for years until the market is fully saturated.

Further, even if one accepts that Autodesk should be allowed to prevent a user who purchases a new version at a lower price from reselling the older version, it is not self-evident that it should have recourse to anything more than a claim for breach of contract against the user. In other words, whatever efficiency restricting the immediate user achieves does not automatically justify an ability to control all downstream and future resale and does not immediately translate into a property right against any subsequent user. Even if the existence of a secondary market is problematic, and suing secondary market dealers might be more effective than suing Autodesk's own customers,²⁷⁸ the problem is short term, mainly during the transition from one version to a newer one. Given that permitting Autodesk to design its initial transaction in a way that allows it to control any subsequent transaction can have long-term

278. Interestingly, the decision in *Vernor* indicates that Autodesk sued not only Vernor but also CTA, the customer who sold the copies to Vernor. The parties stipulated to entry of a permanent injunction against CTA from directly or contributorily infringing Autodesk's copyrights. There is no mention of any damage award for either breach of contract or copyright infringement, suggesting that proving damages had not been easy.

and unknown external costs that neither Autodesk nor its immediate user internalize, there is a real possibility that a remedy for infringement exceeds the problem requiring solution.

The outcome in *Vernor* resulted from the court's literal reading of § 109(a), which led it to conclude that the first sale doctrine did not apply because the buyer was not an owner of the copy but merely a licensee and "[i]n its current form, [the first sale doctrine] allows the 'owner of a particular copy' of a copyrighted work to sell or dispose of his copy without the copyright owner's authorization."²⁷⁹ That is, for the Ninth Circuit § 109(a) was the exclusive basis for the first sale doctrine, and therefore any policy consideration should be left to Congress.²⁸⁰

Prior to *Kirtsaeng*, this position may be regarded as defensible by showing deference to Congress and expressing judicial restraint. Post-*Kirtsaeng*, however, this position may be viewed as a palpable error. If, as both the majority and the dissent in *Kirtsaeng* held, the first sale doctrine has a common law component that is not limited by the wording of § 109(a), it may be the duty of a court to apply the doctrine in a common law fashion and refuse to do that only "when a statutory purpose to the contrary is evident."²⁸¹ In that particular case, Autodesk would bear the burden of showing that Congress had intended to permit it to impose a nonexhausting restraint on its software products for the duration of the copyright.

D. UMG Recordings, Inc. v. Augusto

The last of the recent Ninth Circuit cases involving the first sale doctrine is *UMG v. Augusto*.²⁸² UMG is one of the world's largest music companies.²⁸³ Like many music companies, UMG ships specially produced promotional CDs to a large group of industry insiders, such as music critics and radio programmers.²⁸⁴ Many of the promotional CDs bear a statement such as the following:

279. *Vernor*, 621 F.3d at 1107.

280. *Id.* at 1115.

281. *Kirtsaeng v. John Wiley & Sons, Inc.*, 133 S. Ct. 1351, 1363 (2013).

282. *UMG Recordings, Inc. v. Augusto*, 628 F.3d 1175 (9th Cir. 2011).

283. *Id.* at 1177.

284. *Id.*

This CD is the property of the record company and is licensed to the intended recipient for personal use only. Acceptance of this CD shall constitute an agreement to comply with the terms of the license. Resale or transfer of possession is not allowed and may be punishable under federal and state laws.²⁸⁵

Troy Augusto was not a recipient of such CDs but was able to obtain numerous promotional CDs that he later sold on eBay.²⁸⁶ UMG brought an action against Augusto, claiming that selling the CDs infringed on its distribution right. The Ninth Circuit affirmed the district court's grant of summary judgment in favor of Augusto.²⁸⁷ The court held that the first sale doctrine applied in these circumstances.²⁸⁸ In other words, the distribution of the promotional CDs constituted a sale within the meaning of § 109(a) of the Copyright Act. As a result, the recipients became owners of the copies and were free to dispose of them as they saw fit, and the subsequent sale by Augusto did not infringe on UMG's copyright.²⁸⁹ In reaching this conclusion, the court distinguished *Vernor*, finding that although UMG placed written restrictions on the labels of the CDs, it had not established that the restrictions on the CDs created a license agreement rather than a sale.²⁹⁰

Unlike *Costco*, *Kirtsaeng*, or *Vernor*, *UMG v. Augusto* does not seem to be a case of price discrimination. The promotional CDs are sent to industry insiders before the CDs are released for purposes of promoting and advertising the release of the new CD. The promotional CDs are similar but not equal to the new CDs, as they may contain fewer songs and may not include the artwork included with the new CDs.²⁹¹

Augusto is a case that gets very close to being one that may actually justify the restraints that UMG attempted to impose on the

285. *Id.*

286. *Id.*

287. *Id.* at 1183.

288. *Id.* at 1177.

289. *Id.* at 1180.

290. *Id.* at 1181. The court also found that the recipients were free to dispose of them as they saw fit under the Unordered Merchandise Statute, 39 U.S.C. § 3009.

291. *UMG Recordings, Inc. v. Augusto*, 558 F. Supp. 2d 1055, 1058 (C.D. Cal. 2008), *aff'd*, 628 F.3d 1175 (9th Cir. 2011).

recipients of the promotional CDs. The key difference between *Augusto* and a case like *Vernor* is that the promotional CDs are sent *before* the new CDs are released. The difference is crucial because the insiders receiving the CDs are not ordinary consumers. In fact, those insiders can be regarded as coproducers or collaborators, at least to some extent.

Playing songs remains the most effective means for promoting new musical artists, and radio stations play a crucial role in creating and shaping the demand for new songs.²⁹² The relationships between record companies and insiders who receive the promotional CDs are symbiotic. Feedback received from insiders can be important for making any last decisions before the release of a new CD, and the timing and manner in which the promotional CDs are distributed provide important signals to insiders about the promotional efforts the record company intends to exert. These signals help radio stations better forecast the changing tastes of their listeners.²⁹³ But sending the promotional CDs involves risks. For example, premature leaks may undermine carefully planned scheduling surrounding the release of new CDs or may release materials that might not ultimately be included in the released disk.²⁹⁴ All of this suggests that there are legitimate reasons for restraining the free distribution of promotional CDs, at least temporarily.

But the temporal issue is crucial.²⁹⁵ Even if contractual obligations may not be sufficient to guarantee optimal control when it is necessary, UMG's notice and legal theory is overkill. The notice on the promotional CDs indiscriminately prohibited any sale or transfer of possession, even transfers occurring long after the reasons justifying the restrictions no longer exist, and even by people who cannot be regarded as coproducers in any way. The promotional CDs are valuable for some collectors, and they may be valuable in

292. Ariel Katz, *The Potential Demise of Another Natural Monopoly: Rethinking the Collective Administration of Performing Rights*, 1 J. COMPETITION L. & ECON. 541, 583 (2005); Brief for Amicus Curiae Recording Industry Association of America in Support of Reversal at 7, *UMG Recordings*, 628 F.3d 1175 (No. 08-55998) [hereinafter RIAA Brief].

293. PETER M. THALL, *WHAT THEY'LL NEVER TELL YOU ABOUT THE MUSIC BUSINESS* 135 (2006).

294. RIAA Brief, *supra* note 292, at 10.

295. *Cf.* Katz, *Substitution*, *supra* note 151, at 140–47.

the future for researchers, or for other purposes that is presently unknown and unknowable. The first sale doctrine permits the free circulation of such CDs and thereby contributes to the preservation of such promotional CDs, which may be socially valuable in the long run.

E. Capitol Records v. ReDigi

*Capitol Records v. ReDigi*²⁹⁶ dealt with the application of the first sale doctrine to works in digital form. Capitol Records brought action against ReDigi, a website operator, alleging that ReDigi's service, which allowed individuals to resell digital music files that were lawfully purchased on iTunes, amounted to copyright infringement. The district court described the dispute as "a fundamental clash over culture, policy, and copyright law," but declared that "[b]ecause this is a court of law and not a congressional subcommittee or technology blog, the issues are narrow, technical, and purely legal."²⁹⁷ It ruled in favor of the record company, while rejecting ReDigi's legal and policy arguments, noting that it was left for Congress to deem the limitations on the applicability of the first sale doctrine to digital goods outmoded, not for the court.²⁹⁸

ReDigi's service permitted individuals who purchased songs on iTunes to resell them, describing itself as "a 'virtual' marketplace for 'pre-owned' digital music."²⁹⁹ ReDigi attempted to configure its service in such a way that the sale of the music files would not involve multiplication of copies. It asserted that its process "involve[d] 'migrating' a user's file, packet by packet—'analogous to a train'—from the user's computer to the Cloud Locker so that data does not exist in two places at any one time."³⁰⁰ It hoped that this way, its activities would be considered a transfer of possession of copies—permitted under the first sale doctrine—rather than a reproduction of the work, which the first sale doctrine does not

296. *Capitol Records, LLC v. ReDigi Inc.*, 934 F. Supp. 2d 640 (S.D.N.Y. 2013).

297. *Id.* at 645.

298. *Id.* at 656.

299. *Id.* at 645.

300. *Id.*

allow. The court disagreed. It held that “the plain text of the Copyright Act makes clear that reproduction occurs when a copyrighted work is fixed in a new *material object*.”³⁰¹ In the court’s view, the crucial feature of a reproduction is the fixation of the work in a new material object, whether or not multiplication of copies exists.³⁰²

This holding proved fatal to ReDigi first sale defense, because according to the court, the first sale doctrine was “limited to assertions of the *distribution right*. [But] [b]ecause the Court has concluded that ReDigi’s service violates Capitol’s reproduction right, the first sale defense does not apply to ReDigi’s infringement of those rights.”³⁰³ The court also found § 109(a) was inapplicable because “as an unlawful reproduction, a digital music file sold on ReDigi is not ‘lawfully made under this title.’”³⁰⁴ Finally, “the statute protects only distribution by ‘the owner of a *particular* copy or phonorecord . . . of *that* copy or phonorecord. . . . [but] [h]ere, a ReDigi user owns the phonorecord that was created when she purchased and downloaded a song from iTunes to her hard disk. But to sell that song on ReDigi, she must produce a new phonorecord on the ReDigi server. Because it is therefore impossible for the user to sell her ‘particular’ phonorecord on ReDigi, the first sale statute cannot provide a defense.”³⁰⁵ In the court’s view, “the first sale defense is limited to material items, like records, that the copyright owner put into the stream of commerce.”

While the court gave a nod to ReDigi’s “attractive policy argument,” namely the argument that “refusal to apply the first sale doctrine to its service would grant Capitol “a Court sanctioned extension of rights under the [C]opyright [A]ct . . . which is against policy, and should not be endorsed by this Court,” the court dismissed them as irrelevant in light of Congress’s choice to limit the first sale doctrine to the lawful owner’s “particular” phonorecord. Accepting ReDigi’s argument would, in the eyes of the court, amount to an amendment of the Copyright Act, which “is a

301. *Id.* at 648.

302. *Id.* at 650.

303. *Id.* at 655 (citations omitted).

304. *Id.*

305. *Id.*

legislative prerogative that courts are unauthorized and ill-suited to attempt.”³⁰⁶

Nevertheless, the court did not shy away from relying on policy argument to reject ReDigi’s arguments and endorse Capitol’s. When rejecting the argument that its reading of § 109(a) “would in effect exclude digital works from the meaning of the statute,”³⁰⁷ the court explained that § 109(a) “still protects a lawful owner’s sale of her ‘particular’ phonorecord, be it a computer hard disk, iPod, or other memory device onto which the file was originally downloaded.” The court acknowledged that this view of the first sale doctrine “clearly presents obstacles to resale that are different from, and perhaps even more onerous than, those involved in the resale of CDs and cassettes,” but noted that “the limitation is hardly absurd—the first sale doctrine was enacted in a world where the ease and speed of data transfer could not have been imagined.”³⁰⁸ In reaching this conclusion, the court relied on a 2001 United States Copyright Office Report stating that the first sale doctrine should not apply to the distribution of digital works because “[t]he ability of such ‘used’ copies to compete for market share with new copies is thus far greater in the digital world.”³⁰⁹ The reason for that being that unlike physical copies of works which degrade with time and use, digital information does not degrade, and “[t]he ‘used’ copy is just as desirable as (in fact, is indistinguishable from) a new copy of the same work.”³¹⁰ And since “[t]ime, space, effort and cost no longer act as barriers to the movement of [digital] copies . . . [the] natural brake on the effect of resales on the copyright owner’s market, no longer exists in the realm of digital transmissions.”³¹¹

Like many of the arguments favoring no, or a very weak, first sale doctrine, this argument focuses on the short-term benefits that may accrue to copyright owners if the doctrine could be eliminated, while ignoring the resulting the long-term harms to the public. Even if a

306. *Id.*

307. *Id.* at 656.

308. *Id.*

309. *Id.* (quoting U.S. COPYRIGHT OFFICE, LIBRARY OF CONGRESS, DIGITAL MILLENNIUM COPYRIGHT ACT REPORT, § 104 (2001) [hereinafter DMCA Report]).

310. *Id.* (quoting DMCA Report, *supra* note 310).

311. *Id.*

“used” copy is functionally indistinguishable from the “original” copy, it is not clear why it is necessarily a public policy problem that requires a legal solution.³¹² It can be a problem for the copyright owner who wants to price discriminate, and the availability of cheaper “used” copies undermines its ability to do that, but at the same time, the ability to resell copies may increase some buyers’ willingness to pay a higher price for the original in the first place. Importantly, the Copyright Office’s argument ignores that unlike the case of unauthorized reproduction that increases the supply of copies, the first sale does not change the quantum of copies that are circulating: “used” copies are available only because the copyright owner has already sold them in the first place. This is not to say that applying the first sale doctrine to digital works may not pose challenges to copyright owners, but it is far from obvious that those challenges are insurmountable to the point of justifying the doctrine’s abolition. Whatever those challenges are, they may tend to concentrate during or around the time of production and initial distribution and diminish over time, as described above. Like all other reasons that may justify post-sale restraints, the challenges arising from the distribution of digital works may justify some short-term contractual or organizational solutions, but it is far from obvious that they justify burdening digital works with restraints on their alienation lasting the entire duration of the copyright.

F. Summary

Like the Ninth Circuit in *Vernor* and *Omega*, *ReDigi* relied exclusively on the statutory language of § 109(a).³¹³ This is regrettable. *ReDigi* was decided on March 30, 2013, merely eleven days after *Kirtsaeng*, and even though the ruling cites *Kirtsaeng*, it ignored what might be the most important lesson from that case, namely that the first sale doctrine’s common law roots and the policy that underlies it may compel its application beyond the scope of § 109(a).

312. It should be noted that a “used” digital copy is not always a functional equivalent to an original. For example, books, movies, or music albums, may serve as popular gifts. But giving another person a “used” copy may look cheap and undermine some of the social significance of the act of gift giving.

313. *See supra*, text accompanying note 305.

In fact, had the court in *ReDigi* appreciated the full meaning of *Kirtsaeng*, it would have been open to decide the case differently. Justice Breyer's invocation of Lord Coke's seventeenth century writing, rather than being arcane, invites, or indeed compels, a refreshing appreciation of the first sale doctrine, including in its application to digital goods. Lord Coke's words, as cited by Justice Breyer, are worth reproducing here:

[If] a man be possessed of . . . a horse, or of any other chattell . . . and give or sell his whole interest . . . therein upon condition that the Donee or Vendee shall not alien[ate] the same, the [condition] is voi[d], because his whole interest . . . is out of him, so as he hath no possibilit[y] of a Reverter, and it is against Trade and Traffi[c], and bargaining and contracting betwee[n] man and man: and it is within the reason of our Author that it should ouster him of all power given to him.³¹⁴

The ruling in *ReDigi* is preoccupied with “thingness.” Its holding that the reproduction right had been infringed was “of course, confirmed by the laws of physics. It is simply impossible that the same ‘material object’ can be transferred over the Internet.”³¹⁵ Similarly, the first sale doctrine applies only to the transfer of the “particular” copy or phonorecord. Thus, it “still protects a lawful owner’s sale of her ‘particular’ phonorecord, be it a computer hard disk, iPod, or other memory device onto which the file was originally downloaded,” but not the transfer of the digital file apart from the physical thing in which it is embedded. Thus, it ignores one of the most important insights of modern legal thought, namely that “the institution of property is not concerned with scarce resources themselves (‘things’), but rather with the rights of persons with respect to such resources,”³¹⁶ and it ignores that having copyright in a work means that the copyright owner has “has certain rights with respect to reproduction [and specific other uses] of [that work], and that these rights are separate and distinct from the rights that exist with respect to particular physical copies of the [work].”³¹⁷

314. *Kirtsaeng v. John Wiley & Sons, Inc.*, 133 S. Ct. 1351, 1363 (2013).

315. *ReDigi*, 934 F. Supp. 2d at 649.

316. Thomas W. Merrill, *Property and the Right to Exclude*, 77 NEB. L. REV. 730, 731–32 (1998).

317. *Id.* at 732.

What Justice Breyer, via the quote from Coke, reminds us is that the legal significance of a sale of a horse, a book, or any other chattel lies not in the transfer of the horse or the book, but in the transfer of the seller's interest in them: it is about the transfer of one's legal interest with respect to the horse or the book, rather than the transfer of the horse or the book itself.

Thus, when a person buys a song from iTunes, what the person receives, from a legal standpoint, is not a "thing" but permission to do certain acts that prior to the transaction it did not have the right to do, for example, making a reproduction of the digital file on the buyer's computer and on five additional devices.³¹⁸ A digital first sale doctrine, therefore, means that the buyer transfers her use privilege to another person, and thereby deprives herself thereof. Like a reseller of a book, who transfers her interest in the particular copy to the buyer, the reseller of a digital song or a digital book, transfers her use privileges to another. Focusing on the transfer of one's *rights* with respect to a resource, be it a horse, book, CD, iTunes song, or e-book, rather than the transfer of the resource-thing, renders the difference between the book and the ebook meaningless. The court in *ReDigi* might have felt constrained by the statutory language of § 109(a), but had it appreciated the deeper meaning of the holding in *Kirtsaeng*, it might have realized that what *ReDigi* did was helping individuals who wanted to transfer the set of privileges that they had lawfully obtained (permission to download a song, make a specified number of additional reproductions, etc.) to another person. Those individuals did not buy a 'particular' copy, but paid for, and consequently owned, a set of privileges to make certain non-particular copies. *ReDigi* helped them transfer that set of privileges to another person. As codified, the first sale doctrine may refer to the transfer of particular owned copies, but the common law principle that the Congress codified but not abrogated is not so limited.

It may be argued that a digital first sale may still go the way of the dodo because copyright owners (or their authorized resellers)

318. iTunes Store, *Terms and Conditions*, APPLE, available at <http://www.apple.com/legal/internet-services/itunes/us/terms.html#SERVICE> ("Usage Rules" (i) You shall be authorized to use iTunes Products only for personal, noncommercial use. (ii) You shall be authorized to use iTunes Products on five iTunes-authorized devices at any time, except for Content Rentals (see below).").

may characterize their transactions as licensing transactions rather than sales, and further stipulate that the license is nontransferable.³¹⁹ But the question is not whether they would attempt to do so, but whether the law ought to uphold those attempts. Surely, the seventeenth century man who possessed the horse and gave or sold his whole interest therein upon condition that the Donee or Vendee shall not alienate the horse, preferred giving or selling an encumbered horse over giving or selling a nonencumbered one. And presumably, the person who accepted receiving the encumbered horse might have found the condition acceptable or even attractive (e.g., if the price were lower). Similarly, the Bobbs-Merrill Company preferred to sell books and maintain their resale prices downstream, some booksellers were happy to accept those terms, and John Wiley & Sons preferred that the books that it sold in Thailand would not be imported to the United States.

However, notwithstanding the desire of the owner of the horse or the copyright owner in the book to impose such post-sale restraints on alienation, Lord Coke, Justice Day and Justice Breyer have all declined to include such power in the owners' bundle of rights, because recognizing such power might lead to parades of horrors, or externalities, which the owners—as well as their immediate transacting parties—would rationally ignore. Recognizing such power, in Coke's words, would be “against Trade and Traffi[c], and bargaining and contracting.” As Justice Breyer held, “[a] law that permits a copyright holder to control the resale or other disposition of a chattel once sold is similarly ‘against Trade and Traffi[c], and bargaining and contracting’ . . . [because] [w]ith these last few words, Coke emphasizes the importance of leaving buyers of goods free to compete with each other when reselling or otherwise disposing of those goods.”³²⁰ This rationale holds true to digital goods as well. Therefore, the fact that copyright owners (or their authorized resellers) might prefer characterizing their digital transactions as the grant of nontransferable licenses in order to impose post-sale restraints does not determine the fate of the first sale doctrine. The question is whether the law, which has consciously

319. *E.g.*, *Vernor v. Autodesk, Inc.*, 621 F.3d 1102, 1111 (9th Cir. 2010).

320. *Kirtsaeng v. John Wiley & Sons, Inc.*, 133 S. Ct. 1351, 1363 (2013).

and consistently declined to give them such power, ought to change its mind with respect to digital copies.

In all of the cases discussed in this Part, copyright owners might have had plausibly legitimate reasons to be interested in imposing post-sale restraints. I say “plausibly legitimate” reasons to indicate that the restraint seems to have been designed to advance a goal that is not plainly anticompetitive and might even be procompetitive, and to distinguish those cases from a case such as *Bobbs-Merrill* in which the restraint was the means to enforce an industry-wide cartel.³²¹ From an antitrust perspective, the distinction is crucial: while modern antitrust law may treat the latter as per se illegal, not only will it not necessarily condemn the former type of restraints, it might even embrace them. But the insight that post-sale restraints may be efficient, an insight that has been crucial to the maturity of antitrust law, does not automatically translate into a pseudo-equivalent IP rule that allows IP owners to adopt post-sale restraints in a manner that would be binding everyone.

All that the insights from modern antitrust teach is that when firms jointly participate in a productive enterprise that is prone to opportunism, various enforceable restrictions may be necessary, as stated in Proposition 2, which is why *agreements* implementing such restraints are no longer treated as per se unlawful. But as Proposition 3 states, the economics of joint production do not normally justify enforceable long-lasting post-sale restraints or restraints imposed on third parties. Therefore, enforcing post-sale restraints should not normally be part of the property bundle. Indeed, it is far from obvious that the benefits of those restraints could not be mostly achieved through other means, and in any event it is hard to see how any of the claimed benefits and proffered justifications in any of these cases would justify burdening the works with nonexhausted restraints, enforceable in rem, for the entire duration of the copyright term. Since copyright owners and their immediate transacting parties do not internalize the social harms that such long-lasting restraints impose, the “parade of horrors” that they may create is not a speculative outcome, but a probable one.

321. See *supra*, Part III.A.

Therefore, as stated in Proposition 4, working around the first sale doctrine should be presumptively invalid. Courts should refuse to enforce license conditions or contract terms limiting the ability of the user to resell goods embodying IP rights unless the IP owner can demonstrate that the restraint is necessary and superior to other means to achieve efficiency.

Lord Coke surely had not read Coase, but his conclusion that restraints on alienation are “against Trade and Traffi[c], and bargaining and contracting betwee[n] man and man”³²² is no less wise.

X. CONCLUSION

Despite over a hundred years of adjudication, courts have never been able to draw the exact contours of the first sale doctrine or fully articulate its rationale. Recently, insights borrowed from modern antitrust law and economics have been invoked to provide a seemingly robust theoretical foundation for undermining exhaustion rules or narrowing their scope, thereby strengthening IP owners’ control over downstream distribution and use of the goods they produce. It has been suggested that just as antitrust law has recognized the efficiency of post-sale restraints and relaxed its hostility toward them, so should IP law permit their imposition and provide remedies for their breach. This Article shows that, with the exception of certain instances, this trend is misguided and should be resisted, not because the insights from modern antitrust are irrelevant, but because insights from modern antitrust do not support the case against the first sale doctrine. The main benefits of post-sale restraints involve situations of imperfect vertical integration between coproducing or collaborating firms, and those benefits occur during the production and distribution phases or shortly thereafter. In such situations, contracting around the first sale doctrine should be permitted, and agreements containing such restraints should not be automatically condemned. Beyond such limited circumstances, however, the first sale doctrine promotes important social and economic goals: it promotes efficient use of

322. *Kirtsaeng*, 133 S. Ct. at 1363.

goods embodying IP, guarantees their preservation, and facilitates user innovation, while minimizing transaction costs that otherwise might impede those goals. When a closer look is taken at what modern antitrust law can teach, it can be seen that it confirms the validity and supports the continued vitality of the first sale doctrine.