Demystifying Patent Holdup

Thomas F. Cotter
University of Minnesota Law School, cotte034@umn.edu

Erik Hovenkamp
USC Gould School of Law, ehovenkamp@law.usc.edu

Norman Siebrasse
University of New Brunswick, norman.siebrasse@unb.ca

Follow this and additional works at: https://scholarlycommons.law.wlu.edu/wlulr

Part of the Intellectual Property Law Commons, and the Law and Economics Commons

Recommended Citation

Demystifying Patent Holdup

Thomas F. Cotter[†]

Erik Hovenkamp[‡]

Norman Siebrasse[*]

Abstract

Patent holdup can arise when circumstances enable a patent owner to extract a larger royalty ex post than it could have obtained in an arms length transaction ex ante. While the concept of patent holdup is familiar to scholars and practitioners—particularly in the context of standard-essential patent (SEP) disputes—the economic details are frequently misunderstood. For example, the popular assumption that switching costs (those required to switch from the infringing technology to an alternative) necessarily contribute to holdup is false in general, and will tend to overstate the potential for extracting excessive royalties. On the other hand, some commentaries mistakenly presume that large fixed costs are an essential ingredient of patent holdup, which understates the scope of the problem.

In this Article, we clarify and distinguish the most basic economic factors that contribute to patent holdup. This casts light on various points of confusion arising in many commentaries on the subject. Path dependence—which can act to inflate the value of a

---

technology simply because it was adopted first—is a useful concept for understanding the problem. In particular, patent holdup can be viewed as opportunistic exploitation of path dependence effects serving to inflate the value of a patented technology (relative to the alternatives) after it is adopted. This clarifies that factors contributing to holdup are not static, but rather consist in changes in economic circumstances over time. By breaking down the problem into its most basic parts, our analysis provides a useful blueprint for applying patent holdup theory in complex cases.

Table of Contents

I. Introduction

For more than ten years now, patent infringement litigation filed by owners of standard-essential patents (SEPs)[1] and by patent assertion entities (PAEs)[2] has caused courts and other decision makers throughout the world to reconsider the conventional practice of awarding the prevailing patent owner a permanent injunction (as opposed to an ongoing royalty for prospective infringing sales). In the United States, for example, the Supreme Court in its 2006 decision in *eBay Inc. v. MercExchange, L.L.C.*[3] rejected the long-held view that patent owners are entitled to injunctions almost as a matter of right, and ruled instead that courts should consider four factors (irreparable harm, inadequacy of legal remedies, the balance of hardships, and the public interest) to determine whether injunctive relief is appropriate.[4]

As a result, U.S. courts today rarely award injunctions to SEP owners and PAEs.[5] Meanwhile, courts in other countries have also made it more difficult for SEP owners to obtain injunctions,

---

1. SEPs are patents that:

> read on aspects of technical standards, like the wireless communications standards that are adopted by standard setting organizations (SSOs) such as the Institute of Electrical and Electronics Engineers (IEEE), the International Telecommunications Union (ITU), and the European Telecommunications Standards Institute (ETSI). Often these organizations require their members to disclose any patents they own that read on any aspect of a standard that the organization is considering or has adopted, and to commit to licensing those patents to third parties on fair, reasonable, and nondiscriminatory (FRAND) terms.

THOMAS F. COTTER, PATENT WARS: HOW PATENTS IMPACT OUR DAILY LIVES 230 (2018).

2. PAEs, sometimes referred to pejoratively as "patent trolls," are entities that "acquire patents from third parties and seek to generate revenue by asserting them against alleged infringers" which "already use . . . the patented technology." FED. TRADE COMM'N, PATENT ASSERTION ENTITY ACTIVITY: AN FTC STUDY 1 (2016) [hereinafter FTC STUDY], https://perma.cc/YK9Q-TU8V (PDF).

3. 547 U.S. 388 (2006).

4. *Id.* at 391–92.

5. For a survey of the relevant empirical studies and case law, see Thomas F. Cotter & John M. Golden, *Empirical Studies Relating to Patents—Remedies, in* 2 RESEARCH HANDBOOK ON THE ECONOMICS OF INTELLECTUAL PROPERTY LAW 390, 398–403 (Peter Menell et al. eds., 2019).

although the standards used to decide whether an injunction is warranted in an SEP case—grounded variously in antitrust law, contract law, the law of patent remedies, or the civil law doctrine of abuse of right—are hardly uniform.[6] Moreover, courts outside the United States, for the most part, have continued routinely granting injunctions in other patent disputes;[7] a practice that leads some observers to worry that, just when trolling behavior has started to decline in the United States, countries such as Germany and China (or possibly Europe's soon-to-be-up-and-running Unified Patent Court) will become troll magnets.[8]

On the other side of the coin, some analysts argue that the perceived abuses that have convinced U.S. authorities to deny injunctions in certain types of cases are a myth, and that it's U.S. practice that has taken a wrong turn post-*eBay*. Just this past year, for example, the newly appointed head of the U.S. Department of Justice's Antitrust Division has made a series of speeches critiquing courts' reluctance to grant injunctions in SEP

---

6.   For discussion of the foreign case law, see Jorge L. Contreras et al., *The Effect of FRAND Commitments on Patent Remedies*, *in* PATENT REMEDIES AND COMPLEX PRODUCTS: TOWARDS A GLOBAL CONSENSUS 160, 175–201 (Jorge Contreras et al. eds., 2019) [hereinafter COMPLEX PRODUCTS].

7.   *See* Norman V. Siebrasse et al., *Injunctive Relief*, *in* COMPLEX PRODUCTS, *supra* note 6, at 115, 141 ("As a broad generalization, countries with civil law systems tend to award injunctive relief to a prevailing patentee routinely, and in some countries, effectively as a matter of right . . . .").

8.   *See, e.g.*, INTELLECTUAL PROPERTY 2 INNOVATE, SUPPORTING INNOVATION IN EUROPE THROUGH A BALANCED PATENT SYSTEM: A PAPER RESPONDING TO THE EUROPEAN COMMISSION'S IP PACKAGE 8 (2018), https://perma.cc/6MAX-7F5N (PDF):

> [T]he ability to threaten to demand an automatic injunction . . . is a weapon leveraged by PAEs even though their true goal is to receive payment, not block products from the market. That weapon will become all the more powerful when the UPC [Unified Patent Court] begins operations and an injunction across all Member States is possible.

See also Kaya Elkiner & Kyla Payne, *Patent Infringement Claims in China*, LAW SOC'Y GAZETTE (Dec. 23, 2016), https://perma.cc/V4JF-PHMK (last visited Nov. 3, 2019) (noting that the availability of permanent injunctions could result in an increase in patent litigation by non-practicing entities in China) (on file with the Washington and Lee Law Review).

cases,[9] and legislation has been introduced in Congress that, if enacted, would overturn the *eBay* decision.[10]

These differing views over whether, or when, courts should award injunctions in patent disputes often reflect the speaker's understanding of the frequency with which patent owners engage in a practice referred to as "patent holdup." To see why, consider the fact that SEP owners, PAEs, and many other patent owners often are uninterested in using their patent rights to exclude infringers from the market; they are, instead, willing to license these rights for the right price, and a credible threat of obtaining an injunction enhances the owner's ability to negotiate a favorable deal.

Given this reality, many economists and legal scholars have argued that patent owners who can threaten to "hold up" infringers by obtaining injunctions can compel the latter to pay excessive royalties—that is, royalties in excess of what the infringer would have agreed to pay for a license in an arm's length negotiation—in order to avoid abandoning the infringing technology and switching to a noninfringing alternative.[11] In reaching this conclusion, these

9. Makan Delrahim, Assistant Attorney Gen., Antitrust Div., U.S. Dep't of Justice, Keynote Address at the LeadershIP Conference on IP, Antitrust, and Innovation Policy: The Long Run: Maximizing Innovation Incentives Through Advocacy and Enforcement, (Apr. 10, 2018); Makan Delrahim, Assistant Attorney Gen., Antitrust Div., U.S. Dep't of Justice, Keynote Address at the University of Pennsylvania Law School: The "New Madison" Approach to Antitrust and Intellectual Property Law (Mar. 16, 2018); Makan Delrahim, Assistant Attorney Gen., Antitrust Div., U.S. Dep't of Justice, Remarks at USC Gould School of Law—Application of Competition Policy to Technology and IP Licensing: Take It to the Limit: Respecting Innovation Incentives in the Application of Antitrust Law (Nov. 10, 2017).

10. *See* STRONGER Patents Act of 2018, H.R. 5340, 115th Cong. § 106 (2018) (proposing to amend 35 U.S.C. § 283 to state that "[u]pon a finding by a court of infringement of a patent not proven invalid or unenforceable, the court shall presume that (1) further infringement of the patent would cause irreparable injury; and (2) remedies available at law are inadequate to compensate for that injury," effectively overruling *eBay*); STRONGER Patents Act of 2017, S. 1390, 115th Cong. § 106 (2017) (same); Restoring America's Leadership in Innovation Act of 2018, H.R. 6264, 115th Cong. § 12 (2018) (similar); Inventor Protection Act, H.R. 6557, 115th Cong. § 3 (2018) (proposing amendments similar to H.R. 6264, but applicable only to "inventor-owned patents").

11. *See, e.g.*, FED. TRADE COMM'N, THE EVOLVING IP MARKETPLACE: ALIGNING PATENT NOTICE AND REMEDIES WITH COMPETITION 225–35 (2011) [hereinafter FTC REPORT], https://perma.cc/Y9SL-HE5C (PDF) (assessing the extent to which

observers have drawn upon a body of economic literature describing what we will refer to as "classic" holdup, in which one party to an existing contract extracts *ex post* concessions from the other by taking advantage of the latter's investment in transaction-specific investments that cannot readily be deployed for other purposes.[12] If left unchecked, the possibility that such conduct may occur can reduce social welfare by discouraging parties, *ex ante*, from entering into what otherwise could have been mutually advantageous transactions.[13]

---

injunctive relief may give rise to holdup); Thomas F. Cotter, *Patent Holdup, Patent Remedies, and Antitrust Responses*, 34 J. CORP. L. 1151, 1179–80 (2009) (same); Mark A. Lemley & Carl Shapiro, *Patent Holdup and Royalty Stacking*, 85 TEX. L. REV. 1991, 2035–39 (2007) (same); Carl Shapiro, *Injunctions, Hold-Up, and Patent Royalties*, 12 AM. L. & ECON. REV. 280, 302 (2010) (same); *see also infra* Part II.C.

12. *See infra* Part II.B.

13. *See, e.g.*, OLIVER E. WILLIAMSON, THE ECONOMIC INSTITUTIONS OF CAPITALISM: FIRMS, MARKETS, RELATIONAL CONTRACTING 43–68 (1985) (discussing the various economic approaches and behavioral assumptions employed by the "contractual man" when assessing contracts); Victor P. Goldberg, *Regulation and Administered Contracts*, 7 BELL J. ECON. 426, 439 (1976) (detailing "the customers' demand for protection of the right to be served"); Benjamin Klein et al., *Vertical Integration, Appropriable Rents, and the Competitive Contracting Process*, 21 J. L. & ECON. 297 (1978) (discussing "one particular cost of using the market system—the possibility of post-contractual opportunistic behavior"); Oliver Hart, *Incomplete Contracts, in* 2 THE NEW PALGRAVE: A DICTIONARY OF ECONOMICS 752 (John Eatwell et al. eds., 1987); Benjamin Klein, *Why Hold-Ups Occur: The Self-Enforcing Range of Contractual Relationships*, 34 ECON. INQUIRY 444 (1996); Oliver E. Williamson, *The Vertical Integration of Production: Market Failure Considerations*, 61 AM. ECON. REV. 112 (1971). This work in turn built on concepts pioneered by Ronald Coase. *See* R. H. Coase, *The Nature of the Firm*, 4 ECONOMICA 386 (1937). Of course, it has long been observed that someone looking to develop real property owned by multiple separate entities faces a risk that the last owner to agree to terms will "hold out" for a price far in excess of the proportionate value of her plot. Economic analysis of the efficiency consequences of holdout paralleled the development of the holdup literature noted above. *See, e.g.*, Guido Calabresi & A. Douglas Melamed, *Property Rules, Liability Rules, and Inalienability: One View of the Cathedral*, 85 HARV. L. REV. 1089, 1106–10 (1972) (discussing eminent domain); Lloyd Cohen, *Holdouts and Free Riders*, 20 J. LEGAL STUD. 351 (1991) (distinguishing between the concepts of "hold out" and "free rider"); Errol E. Meidinger, *The "Public Uses" of Eminent Domain: History and Policy*, 11 ENVTL. L. 1, 49–50 n. 175 (1980), citing Patricia Munch, *An Economic Analysis of Eminent Domain*, 84 J. POL. ECON. 473 (1976)). Patent law commentators sometimes cite this literature on real estate holdout when discussing the concept of "royalty stacking," which is said to occur when multiple patent owners, acting in their individual self-interest, demand royalties that in the aggregate exceed the social optimum. Cotter, *supra* note 11, at 1163, 1165

Extending this analysis to the patent context, once a manufacturer has committed to using a technology that it discovers, after the fact, is patented, a patent owner armed with an injunction can extract a royalty that reflects not only the technology's economic value—that is, its *ex ante* value in comparison with the next-best available nonpatented alternative[14]—but also some excess above and beyond this, variously described as consisting of the manufacturer's "sunk" costs of adopting the technology[15] and/or the "switching costs" the manufacturer would incur if, *ex post*, it were to abandon the infringing technology and adopt the next-best alternative.[16] Further, if and when patent holdup occurs, it threatens harm not only to short-term (static) efficiency, analogous to the harm caused by classic holdup, but also to the long-term (dynamic) efficiency that the patent system is intended to promote.[17]

Concern about the impact of holdup has been a driving factor behind the aforementioned shift away from granting permanent injunctions in the United States[18] and, to a more limited degree, elsewhere.[19] But if courts are going to withhold injunctions on account of holdup risks, it is crucial that they understand what holdup *is*, and what sort of evidence would be consistent with a

---

n.71. For present purposes, however, we will focus on the economic literature as developed by Williamson and others to address opportunistic behavior in the context of incomplete contracts, since this seems to have the most relevance to the various forms of conduct that have been described as patent holdup.

14. For discussion of the relevant literature, see *infra* note 45 and accompanying text.

15. For the sake of brevity, we will use the term "sunk costs" to refer to the costs the infringer has incurred that are specific to implementing the infringing technology, and cannot be salvaged or redeployed for other purposes. Assets that could be salvaged or redeployed (e.g., for use in implementing the noninfringing alternative) would not factor into our definition of a holdup royalty. *See infra* notes 30–31 and accompanying text (discussing work by Klein et al.); notes 96–99 and accompanying text (presenting our model).

16. *See infra* note 90 and accompanying text.

17. *See infra* note 80 and accompanying text.

18. *See supra* notes 3–5 and accompanying text (describing the decline in issuing permanent injunctions within the United States).

19. *See supra* note 6 and accompanying text (referring to relevant foreign case law).

serious holdup risk in a particular case; but this is often easier said than done, in part due to the literature's ambiguous or inconsistent usage of the term "switching costs." In the general economics literature, a switching cost is said to be "created whenever the consumer makes an investment specific to his current seller that must be duplicated for any new seller."[20]

Applied to the present context, this definition would include as a switching cost the manufacturer's loss of the ability to recoup the sunk costs of adopting the infringing technology.[21] The patent owner's ability to extract some portion of it *ex post* is a source of holdup, albeit one that is not distinct from sunk costs. Perhaps as a result, in the context of patent holdup, observers sometimes use the terms "sunk costs" and "switching costs" almost interchangeably.[22] Other statements, by contrast, appear to distinguish switching from sunk costs by describing the former as including "costs of redesign, investments in additional plant or equipment, any difference in incremental production costs, and any difference in consumers' willingness to pay for the product"[23] (i.e., future costs that would be incurred to adopt an alternative technology) as distinct from the sunk (past) costs of adopting the

---

20.   *See* Paul Klemperer, *Switching Costs*, *in* THE NEW PALGRAVE DICTIONARY OF ECONOMICS, *infra* note 29, at 13385 (defining switching costs); *see also* Joseph Farrell & Paul D. Klemperer, *Coordination and Lock-In: Competition with Switching Costs and Network Effects*, *in* 3 HANDBOOK OF INDUSTRIAL ORGANIZATION 1967, 1977 (Mark Armstrong & Robert Porter eds. 2007) (same). This effect would constitute holdup, however, only if it is not anticipated *ex ante*. To the extent it is anticipated, prices would go up, but they would do so in *ex ante* contracts, not through opportunism.

21.   Viewing this inability to recoup the sunk costs as a loss is not an example of the sunk cost fallacy. Irrecoverable sunk costs are irrelevant to the profitability of an enterprise going forward, but in the present context it is the ability to recover the sunk costs through product sales that the manufacturer forgoes if it abandons the infringing technology. *See* Tun-Jen Chiang, *Trolls and Orphans*, 96 B.U. L. REV. 691, 695 n.10 (2016); Joseph Farrell et al., *Standard Setting, Patents, and Hold-up*, 74 ANTITRUST L.J. 603, 612–13 & nn.35, 37 (2007).

22.   *See* FTC REPORT, *supra* note 11, at 191 ("The ability of patentees to demand and obtain royalty payments based on the switching costs faced by accused infringers, rather than the ex ante value of the patented technology compared to alternatives, is commonly called 'hold-up.'"); *id.* at 191 n.61 ("'Hold-up' is used throughout this report to describe a patentee's ability to extract a higher licensing fee after an accused infringer has sunk costs into implementing the patented technology than the patentee could have obtained at the time of design decisions, when the patented technology competed with alternatives").

23.   *Id.* at 190 n.53.

infringing technology.[24] Moreover, in asserting that switching costs can contribute to holdup, some commentaries might be read as taking the position that redesign and other associated *ex post* costs, as such, are a distinct source of patent holdup.[25] Yet others appear to view switching costs as distinct from sunk costs, but problematic only if the manufacturer has incurred sunk costs.[26]

In this Article, we show that many common descriptions or understandings of the relationship between switching costs and patent holdup can be misleading, if not outright wrong—and that these misunderstandings *matter*, because an incorrect or imprecise view of patent holdup can lead courts either to *underestimate* the risk of holdup under a given set of circumstances (and thus, perhaps, to grant injunctions too readily), or to *overestimate* it in other cases (thus potentially causing them to spot holdup where it does not exist).[27] We believe, however, that if courts and other legal actors are better able to distinguish the conditions that present a serious risk of holdup from those that do not, they will be able to formulate better rules and standards for deciding when to grant or deny injunctions in a wide range of patent disputes including, but certainly not limited to, suits filed by SEP owners and PAEs.

To assist in this endeavor, we present a simple model showing that the key to understanding holdup is path dependence.[28] In general, path dependence refers to any situation in which a decision maker's options (or the relative desirability of such options) have been affected by her prior decisions, causing market conditions to evolve in a particular way. Most applications center on prior decisions that are "self-reinforcing" for one reason or another, making them more likely to be followed in the future, even if better options arise.[29] In the present context, path dependence

---

24. *Id.*

25. *See infra* note 91 and accompanying text.

26. *See infra* note 92 and accompanying text.

27. To be fair, many of our conclusions about the nature of patent holdup can be found, in one form or another, in the existing law-and-economics literature. Misunderstandings nevertheless persist; and, as noted in the text, can result in either over- or underestimates of the risk of holdup in a given case.

28. *See infra* Part III.A.

29. Sunk costs and network effects are prominent sources of path dependence. For example, the imperial measurement system persists in the

means that once an implementer commits to a particular *ex ante* technological choice, the technology's incremental value over alternatives is different *ex post* from what it was *ex ante*. When this happens, an implementer armed with an injunction can extract an *ex post*, holdup royalty up to the sum of (1) the manufacturer's sunk costs of adopting the infringing technology; (2) the difference between the *ex ante* and *ex post* cost of adopting the noninfringing alternative; and (3) the difference between the *ex ante* and *ex post* benefit of adopting the alternative. Put another way, the holdup royalty consists of sunk costs plus the opportunity cost of *not* having chosen the noninfringing alternative *ex ante*.

As a result—and contrary to some common understandings of patent holdup—we show that *ex post* switching costs, in the sense of redesign and other related costs distinct from sunk costs, do not contribute to patent holdup. In addition, although all three of the sources of patent holdup described above share a common origin in path dependence, any one of them can exist independently of the others. Decision makers therefore need to understand not only what the three components are (and what they *are not*), but also to recognize that looking for one of the three sources independently of the others (or, conversely, requiring evidence of all three in any given case) will lead to over- or underestimates of the actual risk of holdup, and thus potentially to injunction standards that are either unduly lax or excessively rigorous.

Part II presents a brief overview of the existing law-and-economics literature on injunctions, as well as both classic and patent holdup. Part III presents our model, along with several paradigm examples to assist legal decision makers in determining whether a particular set of circumstances presents a serious risk of patent holdup, or not. Part IV shows how our analysis of patent holdup holds up, as it were, in light of arguments made by scholars critical of the holdup concept. Part V concludes, followed by a Technical Appendix.

United States and some other nations, despite being inferior to the metric system, simply because it is already entrenched as the standard system and would be cumbersome to switch. For general discussion of path dependence, see, e.g., Steven N. Durlauf, *Path Dependence*, *in* THE NEW PALGRAVE DICTIONARY OF ECONOMICS 10094 (3d ed. 2018); Douglas Puffert, *Path Dependence in Technical Standards*, *in* THE NEW PALGRAVE DICTIONARY OF ECONOMICS 10106 (3d ed. 2018).

II. Injunctions, Classic Holdup, and Patent Holdup

In this Part, we present a brief overview of the economic literature on classic and patent holdup. We also show that the relevant legal literature sometimes omits some important nuances, and that legal analysts therefore may either over- or underestimate the risk of patent holdup under a given set of circumstances.

A. Injunctions

As we noted in the Introduction, until fairly recently, courts in the United States almost always rewarded the prevailing patent owner with an injunction, and courts in other parts of the world still mostly continue to follow this practice.[30] The principal economic argument in favor of awarding injunctions—as opposed to allowing the infringer to continue using the patented invention, subject to an obligation to pay a court-determined ongoing royalty—is derived from Calabresi and Melamed's foundational article *Property Rules, Liability Rules, and Inalienability: One View of the Cathedral*.[31] In brief, the question Calabresi and Melamed addressed is whether it is better to protect an entitlement by means of a "property" rule—which enables the owner of the entitlement to enjoin others from the unauthorized use of the entitlement, and thus compels prospective users to negotiate with the owner for permission to use—or a "liability" rule, under which the user can, in effect, choose to breach and pay court-determined damages.[32]

One advantage of property rules, Calabresi and Melamed observed, is that a threat of injunctive relief channels parties into voluntary negotiations, and thus conserves on adjudication costs.[33] Moreover, the parties themselves often are in a better position than a court would be to accurately determine the value of the entitlement, in which case injunctions can reduce the risk and

---

30. *See supra* notes 5–7 and accompanying text.
31. 85 HARV. L. REV. 1089 (1972).
32. *Id.* at 1092.
33. *Id.* at 1093–98.

attendant costs of judicial error.[34] Alternatively, if there is reason to think that transaction costs or other bargaining obstacles might foil the socially optimal outcome, perhaps because the number of affected parties is large and disperse, or that the error and adjudication costs resulting from judicial determinations would be tolerably low, a liability rule may be preferable.[35]

Analysts inspired by Calabresi and Melamed have often argued that, as a general matter, courts should protect intellectual property (IP) by means of property rules rather than by means of liability rules.[36] The factual underpinning of the argument is that, because IP assets are unique, they are particularly difficult for third parties to value.[37] Injunctions therefore exploit the parties' presumed advantage in predicting the value each could expect to derive from the use of the IP in comparison with other alternatives, and thus in determining the terms, if any, under which licensing would be more valuable to the owner than exclusion.[38]

On this view, then, by forcing the parties to bargain in the shadow of an injunction, a property-rule entitlement conserves judicial resources and reduces the risk of judicial error.[39] This latter advantage in particular looms large if one believes (as some though not all observers do) that judicial errors in the aggregate would not cancel out, but rather that courts would be more likely to undercompensate, rather than overcompensate IP owners, thus potentially undermining the incentive scheme at the heart of patent and other IP laws.[40] To the extent courts employ the legal standards for calculating damages in a manner that is inconsistent or unpredictable, moreover, the resulting uncertainty puts risk-averse parties at a disadvantage and, at the margin, can deter settlements.[41]

---

34. *Id.* at 1124–27.

35. *Id.* at 1105–10.

36. For an overview of the relevant literature, see Cotter & Golden, *supra* note 5, at 392–94.

37. *Id.*

38. *Id.* at 3.

39. *Id.*

40. *Id.* at 12–14.

41. *See* Thomas F. Cotter, *Patent Damages Heuristics*, 25 TEX. INTELL. PROP. L.J. 159, 168 (2017) (contending that a large range of possible outcomes would result in fewer settlements and an increased likelihood of risk-adverse parties

That said, even if one finds the above analysis generally persuasive, there could be advantages to liability rules in specific cases. One might imagine instances, for example, in which the public interest in access to a patented invention (e.g., a life-saving drug) is sufficiently great that a court or agency should sidestep private bargaining and in effect compel the licensing of the invention at a government-determined price. The social cost of granting injunctions, in such cases, arguably outweighs the social benefits (in terms of reducing adjudication and error costs).[42] Indeed, even in the United States prior to *eBay*, courts on rare occasions would decline to enter injunctions in favor of patent owners on public interest grounds.[43]

As suggested in the Introduction, economists and legal scholars today often argue that cases presenting patent holdup risks are yet another example of situations in which the social welfare costs of granting injunctions are potentially very high; high enough, perhaps, to outweigh the adjudication and error costs that may result from denying them.[44] Of course, just because a case presents a risk of patent holdup does not necessarily mean that, if an injunction were to issue, the patent owner would actually demand an excessive royalty; all other things being equal, it may be rational for some patent owners to refrain from extracting holdup rents even when they could do so, in order to avoid inviting retaliation should they find themselves on the receiving end of others' licensing demands.[45]

---

settling on unfavorable terms).

42. On the other hand, it's possible that compulsory licensing in such cases could reduce social welfare if it were to inhibit the incentive to invest in the development of new drugs that would benefit more people over the long term.

43. *See, e.g.*, Cty. of Milwaukee v. Activated Sludge, Inc., 69 F.2d 577, 593 (7th Cir. 1934) (denying injunctive relief when the potential consequence, closing a community's only safe means for disposal of raw sewage, outweighs the interests of the patentee); *see also* Vitamin Technologists, Inc. v. Wis. Alumni Res. Found., 146 F.2d 941, 944 (9th Cir. 1945) (raising, but not deciding, the question of whether an injunction against sales of oleomargarine would be against the public interest).

44. *See supra* note 13 and accompanying text.

45. *See, e.g.*, U.S. DEP'T OF JUSTICE & FED. TRADE COMM'N, ANTITRUST ENFORCEMENT AND INTELLECTUAL PROPERTY RIGHTS: PROMOTING INNOVATION AND COMPETITION 40 (2007) [hereinafter ANTITRUST ENFORCEMENT] (stating that "reputation and business costs" may sometimes mitigate holdup).

Moreover, even if the potential social costs of granting injunctions are higher when a holdup risk is present than they are in other cases, there is no obvious reason to expect the social costs of denying them to be any lower; if anything, these costs could be higher too, if a legal standard that makes injunctions hinge upon a lack of holdup risk itself increases adjudication costs, or if denying injunctions in a wider range of cases invites abuse on the part of accused infringers or other unintended consequences, for reasons we will touch upon in Part III. For these reasons, whether the potential harms of granting injunctions, in cases in which holdup risks are present, outweigh the potential harms of denying them is a matter on which reasonable minds may differ, and with respect to which each legal system must reach its own conclusion.

Be that as it may, a rational assessment of the matter requires a clear understanding of what patent holdup *is*, the harms it threatens, and the conditions under which the risk of encountering it is serious. Armed with this knowledge, courts and other decision makers can then evaluate the evidence and proceed as they see fit. An erroneous understanding of what holdup is, by contrast, risks over- or underestimating the potential harms of granting injunctions in PAE, SEP, and other cases. To better enable policymakers to grasp the issues, therefore, the following sections sketch out the principal insights on holdup, in both the classic and patent sense, as found in the relevant literature, as well as the ambiguities that risk leading decision makers astray. In Part III, we present our own model to isolate the three components of patent holdup as identified, though not always clearly distinguished, in the previous literature. We also offer some advice on how to detect whether holdup risks are present, or not, in a range of real-world cases.

B. Classic Holdup

The classic holdup literature dates back to the 1970s, with work in which Oliver Williamson, Benjamin Klein, and other economists sought to ground various forms of industrial organization in the economics of transaction costs.[46] Williamson in

---

46.   *See supra* note 13 and accompanying text.

particular pioneered the idea that a recurring problem faced by contracting parties is that of opportunism; which Williamson defined as "self-interest seeking with guile."[47] As Williamson observed, "[t]ransactions that are supported by investments in durable, transaction-specific assets experience 'lock-in' effects"—a condition Williamson referred to as "asset specificity"—which makes such investments "risky, in that specialized assets cannot be redeployed without sacrifice of productive value if contracts should be interrupted or prematurely terminated."[48] In principle, parties could minimize such risks if it were possible to address all possible contingencies in advance, but such "exhaustive stipulation, assuming that it is feasible, is itself costly."[49] Contracting parties therefore are faced with a dilemma:

> Although both a have a long-term interest in effecting adaptations of a joint profit-maximizing kind, each also has an interest in appropriating as much of the gain as he can on each occasion to adapt. Efficient adaptations that would otherwise be made thus result in costly haggling or even go unmentioned, lest the gains be dissipated by costly subgoal pursuit. Governance structures that attenuate opportunism and otherwise infuse confidence are evidently needed.[50]

Klein, along with Robert Crawford and Armen Alchian, further developed the theory by showing how a party who is able to renegotiate a contract might be able to extract some portion of the other party's quasi-rents.[51] ("Quasi-rents," otherwise known as variable profits, are equal to total revenue minus total variable or operating costs).[52] To illustrate, Klein et al. used an example involving a printing press owner (Firm A) that agrees to perform printing services for Firm B for a price of $5,500 per day.[53] Firm A's amortized fixed costs are $4,000 per day, the press's salvageable value if it is moved elsewhere is the daily rental

---

47. WILLIAMSON, *supra* note 13, at 47.

48. *Id*. at 52–54.

49. *Id*. at 115.

50. *Id*. at 63.

51. KLEIN ET AL., *supra* note 13, at 298–302.

52. Cheng-chung Lai & Guang-Jong Fann, *Marshall's Quasi-Rent Reconsidered*, 8 HIST. ECON. IDEAS 99, 101 (2000).

53. KLEIN ET AL., *supra* note 13, at 298–99.

equivalent of $1,000, and the daily operating costs are $1,500.[54] Over the long term, Firm A thus expects to break even on its investment in the press, but its (short run) quasi-rent is $3,000 ($5,500 - $1,500 - $1,000).[55] In addition, the example posits that the next-best option for Firm A would have been to contract with Firm C, which offered $3,500 for the same type of services.[56]

Under this contract, A's quasi-rent would have been $1,000 (though it would have lost $2,000 on its investment over time).[57] Now suppose that Firm B could credibly threaten to walk away from the contract unless A agreed to renegotiate the price. Having already incurred the fixed cost of $4,000, A would be better off agreeing to any renegotiated price above $2,000 (leaving it with a quasi-rent of at least $1,000 and a long run amortized loss of at most $2,000), because A's next-best alternative (contracting with C) would leave it even worse off.[58] Put another way, B can appropriate a portion of A's quasi-rent up to the value of B's next-best alternative. Further, if there *were* no alternative customer C *ex post*, B could appropriate up to the entire nonsalvageable fixed cost of $3,000, since having incurred its fixed cost B would be better off accepting this deal than walking away.[59] Alternatively, depending on the facts, it's possible that Firm B could be the party practicing holdup by renegotiating a higher price, if (say) A has incurred sunk costs in reliance on having the printing job completed by a given date.[60]

Of course, parties do not invariably hold one another up in the fashion of this stylized example. First, if either party develops a reputation for sharp practice it may find its opportunities for future transacting with other parties severely limited. Second, even incomplete contracts do not permit B (or A) to credibly demand renegotiation for any reason whatsoever, though as noted above, the cost of drafting tighter contracts, monitoring them, and if necessary, enforcing them is not zero. Third, parties who are

---

54. *Id.*
55. *Id.*
56. *Id.*
57. *Id.*
58. *Id.* at 300.
59. *Id.*
60. *Id*. at 299–300.

aware of the risks and fear being held up can demand a risk premium—that is, a higher (or lower) price that takes into account the perceived risk of future holdup—or decide that the risk is intolerable and not enter into the transaction in the first place. These added costs are a social waste, however, and at the margin could discourage the formation of contracts that otherwise would be in both parties' best interest.

Williamson, Klein, and other holdup theorists therefore predict that private and public actors will invest in reducing the risk of holdup by, for example, vertically integrating different actors along the supply chain or lobbying for government regulation. All such measures come at a cost, however, and none of them are perfect; though over a wide range of conduct they must be reasonably effective, or else systemic holdup would be causing markets to implode left and right.

C. Patent Holdup

To our knowledge, the first authors to apply the holdup concept as described above to patent law issues were Robert Merges and Richard Nelson in their 1990 article *On the Complex Economics of Patent Scope*.[61] In their discussion of the economics of improvement, Merges and Nelson presented an example involving two inventions: an initial invention the value of which over the next-best alternative is $100, and a subsequent invention that incorporates but radically improves upon the original invention, and whose value over alternatives is $900.[62] If the initial invention is patented, the owner can demand a share of the value of the improvement (how much will depend on how good a bargainer she is), thus leaving the second inventor with a

---

61.   Robert Merges & Richard Nelson, *On the Complex Economics of Patent Scope*, 90 COLUM. L. REV. 839 (1990). George Bittlingmayer had earlier invoked the concept of holdout to suggest that patent pools, which license multiple complementary patents owned by different entities, could be viewed as a type of "private condemnation" analogous to eminent domain. *See* George Bittlingmayer, *Property Rights, Progress, and the Aircraft Patent Agreement*, 31 J.L. & ECON. 227, 242 (1988) ("[T]he rationale for a 'private condemnation' of patents that apply to a product is like the rationale for eminent domain.").

62.   Merges & Nelson, *supra* note 61, at 866 n.117.

comparatively small portion of the social value of the improvement.[63] Merges and Nelson viewed this result as analogous to the type of holdup discussed by Klein et al., and argued that it may inefficiently delay introduction of the improvement until the initial patent expires or the parties agree to terms.[64] Suzanne Scotchmer further developed this theme in her 1991 article *Standing on the Shoulders of Giants: Cumulative Research and the Patent Law*, and explicitly noted how the second inventor's incentives are particularly weak if it has incurred sunk costs of research and development prior to negotiating a license with the dominant patent owner.[65]

Carl Shapiro's 2001 book chapter *Navigating the Patent Thicket: Cross Licenses, Patent Pools, and Standard Setting*[66] nevertheless is often seen as the foundational paper on which most of the current discussions of patent holdup are based.[67] Shapiro illustrated the concept of patent holdup by means of an example of a manufacturer who

> could, with ease, invent around a given patent, if that manufacturer were aware of the patent and afforded sufficient lead time. Clearly, in this case the patented technology contributes little if anything to the final product, and any reasonable royalty would be modest at best. But, oh, how the situation changes if the manufacturer has already designed its product and placed it into large scale production before the patent issues . . . . The patentee can credibly seek far greater royalties, very likely backed up with the threat of shutting down the manufacturer if the Court indeed finds the patent valid and infringed and grants injunctive relief. The manufacturer *could* go back and redesign its product, but to do so (a) could well require a major redesign effort and / or cause a significant disruption to production, (b) would still leave potential liability for any products sold after the patent issued before the redesigned products are available for sale, and (c) could present

---

63. *Id.*

64. *Id.* at 865–67.

65. *See* Suzanne Scotchmer, *Standing on the Shoulders of Giants: Cumulative Research and the Patent Law*, 5 J. ECON. PERSP. 29, 32–35 (1991) ("To provide efficient incentives to the second innovator, society should protect the first innovation so narrowly that a new product never infringes and therefore second innovators never have to license.").

66. 1 INNOVATION POLICY AND THE ECONOMY 119, 119–50 (2001).

67. *Id.*

compatibility problems with other products or between different versions of this product. In other words, for all of these reasons, the manufacturer is highly susceptible to holdup by the patentee. I submit that this holdup problem is very real today, and that both patent and antitrust policymakers should regard holdup as a problem of first order significance in the years ahead . . . . The result will be that some companies avoid the mine field altogether, that is, refrain from introducing certain products for fear of holdup. Other companies will lose their corporate legs, that is, will be forced to pay royalties on patents that they could easily have invented around at an earlier stage, had they merely been aware that such a patent either existed or was pending. Of course, ultimately the expected value of these royalties must be reflected in the price of final goods.[68]

Two things are notable in this discussion. First, it implicitly presumes that the correct value of the patented invention is its value prior to the manufacturer's having "designed its product and placed it into large scale production."[69] This is consistent with the mainstream view (which we share) that the correct or baseline royalty for a patented technology is its "*ex ante*" value over the next-best available nonpatented alternative.[70] Second, the

---

68. *Id.* at 125–26.

69. *Id*. at 125.

70. *See* Thomas F. Cotter et al., *Reasonable Royalties* in COMPLEX PRODUCTS, *supra* note 6, at 6, 19, 28:

[There is] a widespread consensus among innovation economists and lawyers that the social value of a technology is its incremental value over the next-best alternative, and that the economic value of a patented technology to an implementer is the (actual or expected) profit or cost saving the implementer derives from the use of the patented technology over the next-best available noninfringing alternative . . . . [and that] the mainstream view in U.S. law [is that the] hypothetical bargain [framework commonly used to estimate reasonable royalties in litigation] should be based only on information that is available to the parties ex ante . . . .

For further discussion of debates concerning the appropriate royalty baseline, see Cotter, *supra* note 11, at 1164–65, 1172–73; Norman V. Siebrasse & Thomas F. Cotter, *A New Framework for Determining Reasonable Royalties in Patent Litigation*, 68 FLA. L. REV. 929, 953 n.73 (2016).We take no position in this Article on the question of whether the patentee's ability to charge a higher *ex post* royalty should be considered holdup if that higher royalty reflects a greater-than-expected advantage of the patented technology over alternatives. In other work, two of us have argued that it should not. *See* Siebrasse & Cotter, *infra* note 82; *cf.* Richard H. Stern, *Who Should Own the Benefits of Standardization and the Value*

discussion identifies three distinct phenomena that could result in the manufacturer agreeing to pay a royalty that exceeds the *ex ante* value of the invention.[71]

The first is the cost that the manufacturer has incurred to "design [ ] its product and place [ ] it into large scale production" prior to being made aware of the existence of the patent.[72] These are *past* costs that were incurred to implement the *current* technology. To the extent these past or "sunk" costs cannot be redeployed to produce a noninfringing product, they are, to use Williamson's terminology, "asset-specific" and therefore capable of generating lock-in effects.[73] Much of the subsequent literature on patent holdup refers to these investments simply as "sunk costs," and identifies sunk costs as a principal source of patent holdup.[74]

It Creates?, 19 MINN. J.L. SCI. & TECH. 107, 125 (2018). The present Article nevertheless concentrates on the three principal sources of patent holdup as identified in the literature, all of which share a common source in path dependence, and it remains agnostic on whether other conduct on the part of patent owners should or should not be labeled "holdup."

71. *See* Shapiro, *supra* note 66, at 125–26 ("[F]or all of these reasons, the manufacturer is highly susceptible to holdup by the patentee.").

72. *See id.* at 125. As noted in the text above, these costs are commonly referred to in the patent holdup literature as sunk costs. That literature typically does not make explicit reference to the "potential liability for any products sold after the patent issued" as a source of patent holdup, but in principle such liability could be viewed as a sunk cost (or perhaps more accurately, sunk liability), to the extent it is an inevitable consequence of the manufacturer's *ex ante* investment in the marketing of an infringing product.

73. Williamson, *supra* note 13, at 52–54.

74. *See, e.g.*, Microsoft Corp. v. Motorola, Inc., No. C10–1823JLR, 2013 WL 2111217, at *10 (W.D. Wash. Apr. 25, 2013), *aff'd*, 795 F.3d 1024 (9th Cir. 2015) ("The threat of hold-up increases as the standard becomes more widely implemented and firms make sunk cost investments . . . ."); Chryssoula Pentheroudakis & Justus A. Baron, *Licensing Terms of Standard Essential Patents: A Comprehensive Analysis of Cases*, JRC SCI. FOR POL'Y REP. 24–25, 138, 149 (2017) (citing Scotchmer, *supra* note 65) (stating that hold-up is intrinsically related to irreversible sunk costs); FTC REPORT, supra note 11, at 5, 8, 10, 26, 50, 92–93, 144, 175, 191 n.61, 215, 222, 225–26, 237 (2011) (discussing how sunk costs lead to patent holdup); ANTITRUST ENFORCEMENT, *supra* note 45, at 35 (describing sunk costs leading to patent hold-up due to increased expenses and low bargaining leverage); Dennis W. Carlton & Allan L. Shampine, *An Economic Interpretation of FRAND*, 9 J. COMPETITION L. & ECON. 531, 535 (2013) (stating that patent owners can demand royalties reflecting "not the incremental value of the patented technology to the licensee . . . , but rather the market power that the holder of [SEPs] has as a result of . . . . sunk investments made by firms implementing that standard"); Bernard Chao, *Horizontal Innovation and Interface Patents*, 2016 WIS. L. REV. 287, 303 (2015) (noting that a company which

The intuition is that if the patent holder can credibly demand to enjoin the use of the infringing technology *ex post*, the manufacturer often would be better off agreeing to pay a royalty that exceeds the technology's *ex ante* value, rather than to write off those asset-specific sunk costs as a loss. The extent of the excess will depend on the parties' relative bargaining power, but as we show in Part III it could be as high as the entire sunk cost.

To be sure, this consequence is not identical to the types of holdup discussed in the classic holdup literature, the archetypal example of which as described above involves one party to an existing transaction demanding some type of *ex post* concession from the other; in the version of "patent holdup" under consideration here, by contrast, there often *is* no initial agreement between a buyer and a seller.[75] (As in the classic holdup scenario, however, one party *is* using its *ex post* leverage to extract a quasi-rent from the other, in this case a royalty that exceeds the

---

has sunk investments into designing, manufacturing and marketing a product is particularly vulnerable to patent assertions); Chiang, *supra* note 21, at 694–96, 698–702; Jorge L. Contreras & Richard J. Gilbert, *A Unified Framework for RAND and Other Reasonable Royalties*, 30 BERKELEY TECH. L.J. 1451, 1487, 1492–93, 1496 (2015) ("[I]rreversible investments can allow a patentee to charge higher royalties than the patentee could obtain in a hypothetical negotiation that occurs before the specific investments are made."); Richard Gilbert, *Deal or No Deal? Licensing Negotiations in Standard-Setting Organizations*, 77 ANTITRUST L.J. 855, 855 (2011) (describing how owners of patents may "hold up" firms or consumers that are "locked-in" to a standard by charging high royalties); Farrell et al., *supra* note 21, at 604, 612–13; Suzanne Michel, *Bargaining for RAND Royalties in the Shadow of Patent Remedies Law*, 77 ANTITRUST L.J. 889, 892 n.9, 892–93, 896, 911 (2011) (defining holdup as the ability to extract higher fees after an accused infringer has sunk costs into implementation); William F. Lee & A. Douglas Melamed, *Breaking the Vicious Cycle of Patent Damages*, 101 CORNELL L. REV. 385, 409–11 (2016) (expanding on the effects of lock-in costs); Jonathan D. Putnam, *Economic Determinations in "FRAND Rate"-Setting: A Guide for the Perplexed*, 41 FORDHAM INT'L L.J. 953, 962 (2018) ("Hold-up is a form of economic opportunism . . . .").

75. In SEP cases, however, some courts understand the SEP owner's commitment to license its patents on FRAND terms to create a contract for the benefit of third parties, the precise terms of which nevertheless are undefined *ex ante. See, e.g.*, Microsoft Corp. v. Motorola, Inc., 795 F.3d 1024, 1030 (9th Cir. 2015) (describing the appellee Microsoft as "a third-party beneficiary to Motorola's RAND commitments"). So at least under these circumstances, the incomplete nature of the contract can enable opportunism, as in the classic holdup scenario.

baseline royalty the parties *would have* agreed to *ex ante*.[76]) In addition, there is not (necessarily) any "guile" on the part of the patent owner, for example in the sense of keeping its patent a secret until the infringer is "locked in"—though guile is surely a *possibility*, and when it is present the patent holder is sometimes accused of engaging in a species of sunk-costs holdup referred to as "patent ambush."[77] (We will discuss this variation on the theme of sunk-costs holdup in Part III.B, but for now it suffices to note that guile is not a necessary condition for the holder to be able to demand an *ex post* royalty that reflects sunk costs).[78] What *is* necessary is that the infringer failed to negotiate a royalty in

---

76.   *See* ANTITRUST ENFORCEMENT, *supra* note 45, at 35 n.11. In this regard, patent holdup may seem more analogous to "holdout" in the real estate context, as described by Williamson, *supra* note 13. In this latter context, there is no preexisting agreement between prospective buyers and sellers, and a prospective seller uses its leverage *ex post* (after other sellers have reached agreement) to extract a disproportionate share of the expected gain.

77.   Cotter, *supra* note 11, at 1188–89.

78.   The absence of guile as a necessary condition for sunk-costs patent holdup leads some commentators to argue that the latter is not a type of holdup at all, but simply a matter of the infringer not wanting to pay what the patent owner demands. Alexander Galetovic & Stephen Haber, *The Fallacies of Patent-Holdup Theory*, 13 J. COMPETITION L. & ECON. 1, 23–24 (2017). The premise, however, that holdup cannot exist in the absence of guile because parties can anticipate, and thus prepare in advance for, the possibility of being held up *ex post* arguably overlooks the potential inefficiencies that can result from having to take such defensive measures *ex ante*. *Id.* In the present context, and for reasons discussed in the text above, even in the absence of guile it is socially inefficient for the patent owner to extract an *ex post* royalty that reflects sunk costs. *Id.* The fact that the inefficiency results from a set of circumstances that differ in some respects from those giving rise to classic holdup does not make it any less inefficient. *Id.*

       Furthermore, even in the classic holdup literature, scholars do not uniformly agree that guile is a necessary condition for holdup. *See* William P. Rogerson, *Contractual Solutions to the Hold-Up Problem*, 59 REV. ECON. STUD. 777, 777 (1992):

   A hold-up problem occurs when two factors are present. First, parties to a future transaction must make non-contractible specific investments prior to the transaction in order to prepare for it. Second, the exact form of the optimal transaction (e.g. how many units if any, what quality level, the time of delivery) cannot be specified with certainty ex ante.

See also Williamson, *supra* note 13, at 47 (stating that "[b]y opportunism, I mean self-interest seeking with guile . . . . More generally, opportunism refers to the incomplete or distorted disclosure of information, especially to calculated efforts to mislead, distort, disguise, obfuscate, or otherwise confuse.").

advance of incurring the sunk cost.[79] Among the possible reasons for not doing so are a lack of actual awareness of the patent or its relevance to the infringer's product, and the perceived need to launch a product quickly before clearing all potential rights (as is said to be common in industries such as telecommunications).[80] Of course, if courts routinely deny injunctions based on the risk of sunk costs holdup, the incentive on the part of implementers to come up with ways of discovering and clearing the necessary rights in advance may be reduced—particularly if the damages courts award in lieu of injunctions are *below* the *ex ante* amount, due to error costs, or are otherwise undercompensatory. This is the so-called holdout phenomenon discussed in Part III below, and we agree with the critics of patent holdup theory that if and when such practices occur the resulting social costs could weigh in favor of granting injunctions, notwithstanding the corresponding risk of sunk costs holdup.[81] For now, though, we note only that, in theory, the risk of being subjected to patent holdup could discourage some implementers from marketing products that might render them vulnerable to being held up *ex post*; and that, as in the classic holdup context, such a risk threatens to make both parties (here, patent owners and implementers) worse off. In addition, in the patent holdup context there is a risk of dynamic inefficiency, in that if the implementer accedes to the owner's demand for a royalty that exceeds the value of the invention over the next-best available nonpatented alternative, the implementer (and, by extension, the consumers who buy the implementer's products) is paying more for the patented invention than it is worth, in terms of its contribution to the state of the art.[82]

---

79. *See* FTC REPORT, *supra* note 11, at 22 (stating that a potential infringer may be able to avoid hold up if royalties are negotiated in advance of incurring sunk costs).

80. *See id.* at 226 (stating the lack of notice of a patent is especially prevalent in IT industries); Lee & Melamed, *supra* note 74, at 404–09 (explaining how the patent system places the practical burden of avoiding infringement on potential infringers); Shapiro, *supra* note 66, at 126 (noting the holdup problem is worse in industries with thousands of patents pending and issued).

81. *See infra* Part III.

82. *See* Norman V. Siebrasse & Thomas F. Cotter, *The Value of the Standard*, 101 MINN. L. REV. 1159, 1184, 1189–90 (2017) (arguing that sunk costs holdup is dynamically inefficient because it results in an excessive incentive to

The second potential source of patent holdup that Shapiro referred to in his 2001 paper is the cost of redesigning a noninfringing product, including the possible disruption of current production.[83] In contrast to the first source of patent holdup, this is a *future* cost of implementing the *alternative* technology.[84] Subsequent literature asserts that this *ex post* cost of implementing a noninfringing alternative will often exceed what the *ex ante* cost of implementing the alternative would have been, and that the manufacturer will agree to pay a royalty exceeding the value of the infringing technology in order to avoid shouldering this cost increase.[85] Note, however, that if the *ex ante* and *ex post* costs of adopting the alternative are identical, the fact that the user would incur these costs *ex post* if it were to switch—and would be willing to pay something to avoid having to do so—does not generate an additional source of holdup rents, since the user would have faced the very same calculus had the parties negotiated *ex ante*.[86]

Third are what Shapiro referred to as "compatibility problems."[87] As developed in the subsequent literature, the intuition here is that, if the manufacturer were to switch to a

---

invent).

83. *See* Shapiro, *supra* note 66, at 125 (stating that redesign "could well require a major re-design effort and/or cause a significant disruption to production").

84. *Id.*

85. *See* ANTITRUST ENFORCEMENT, *supra* note 45, at 35 (explaining the unequal bargaining power in hold-up problems); Pentheroudakis & Baron, *supra* note 74, at 55 (describing the *ex ante* negotiation benchmark); FTC REPORT, *supra* note 11, at 22, 190 (noting that the result of a negotiation can vary significantly depending on when one assumes it occurs); Farrell et al., *supra* note 21, at 650 & n.168 (stating one could examine whether *ex ante* switching costs differ from *ex post* switching costs); Lee & Melamed, *supra* note 74, at 410–11, 410 n.106, 411 n.107 (providing an example in which an infringer is willing to pay costs greater than the cost *ex ante*); Lemley & Shapiro, *supra* note 11, at 2000–05 (explaining how royalty "overcharge," is driven by the threat of obtaining an injunction and the rules by which reasonable royalties are calculated); A. Douglas Melamed & Carl Shapiro, *How Antitrust Law Can Make FRAND Commitments More Effective*, 127 YALE L.J. 2110, 2113 (2018) (explaining that "negotiations between SEP holders and implementers generally take place only after the implementers have used and infringed the technologies claimed by the SEPs"); Shapiro, *supra* note 11, at 286–94.

86. For further explication of this point, see *supra* Part II.A.

87. Shapiro, *supra* note 66, at 125.

noninfringing alternative *ex post*, it might have to either (1) redesign other complementary components or products to interact with that alternative, or (2) forgo the benefits of selling a product that enjoys substantial network externalities (for example, a standard-compliant smartphone that enables consumers to communicate with people who use smartphones made by other manufacturers).[88] In the first case, the "compatibility problems" would simply result in another set of redesign costs, but in the second the harm could be thought of as a loss of expected benefits; and to avoid that loss, the manufacturer might agree, once again, to pay an *ex post* royalty that exceeds the infringing technology's *ex ante* value. This is not an out-of-pocket cost as such, but a future loss of profits associated with switching from the current technology to the alternative. Farrell et al., for example, present a stylized example in which, *ex ante*, a standard setting organization (SSO) chooses between two competing standards, either of which would confer upon consumers a compatibility (network) benefit in the amount of $30.[89] Once the SSO chooses between the two, however, the *ex post* network benefit derived from switching to the alternative is zero:

> If each user's leading alternative to sticking with the standard is unilateral switching, and thus losing compatibility with

---

88. *See* Siebrasse & Cotter, *supra* note 82, at 1172 (citing Michael L. Katz & Carl Shapiro, *Network Externalities, Competition, and Compatibility*, 75 AM. ECON. REV. 424, 424, 434 (1985)):

> Network effects arise when the value a user derives from consumption of a good increases with the number of other agents consuming the good. Communication technologies are a classic example: the more people that have telephones, the more valuable a telephone is to any given person. While network effects may arise in the absence of standards, and de facto standards may arise as a consequence of network effects even in the absence of formal standards, a major reason for the existence of formal standards is to allow the market to coordinate on a single technology in order to reap the benefits of network externalities. Consequently, formal standards are normally associated with network effects.

As we show in Part II, in this context the loss that presents a holdup risk is not the loss of network benefits associated with the infringing (standardized) technology, but (counterintuitively) the loss of network benefits associated with the alternative. *See supra* Part II.B.

89. Farrell et al., *supra* note 21, at 616.

others, then the patent holder's subsequent advantage . . . includes not only its technology's inherent advantage and the value of the user's own sunk investments, but also the value of compatibility to the user . . . . For example, compliance with some telecommunications standards enables a network operator to offer its customers the ability to roam onto compatible networks. In our numerical example, suppose that the two technologies produce incompatible, though otherwise identical, outputs, and that compatibility with other users is worth $30. Then, if all others are expected to stick to the patented technology, adopting it is worth $30 more to each user. The user will adopt the patented technology as long as the royalty demanded is less than $40, composed of $10 of inherent value and $30 of network effect.[90]

The economic literature on patent holdup, as recounted above, has been insightful in identifying these sources of path-dependent patent holdup. As described in the Introduction, however, problems can arise when legal scholars and decision makers try to apply these insights, due to the ambiguous and inconsistent usage of the term—which we ourselves have avoided using in this section up till now—"switching costs." Intuitively, it might seem that "switching costs" are simply the costs that would have to be incurred in switching to the best alternative to the patented technology. To define switching costs in that forward-looking sense, however, neglects the classic holdup problem, which focuses on sunk costs, and sunk costs necessarily are past costs. Accordingly, in one sense of the term, switching costs refer to the user's *ex post* inability to recoup its investment in the infringing

---

90. *Id*. *See also* ANTITRUST ENFORCEMENT, *supra* note 45, at 38 ("Generally, the greater the cost of switching to an alternative standard, the more an IP holder can charge for a license."); FTC REPORT, *supra* note 11, at 192 ("Accused infringers may pay royalties based on the costs of switching to another technology. Switching costs can be prohibitively high when an industry standard is involved."); Lemley & Shapiro, *supra* note 11, at 2016 ("[T]he key point is that it can be extremely costly, or even impossible as a practical matter, to "redesign" a product standard to avoid infringing a patented technology, even if initially an alternative standard could easily have been selected."); Timothy S. Simcoe & Allan L. Shampine, *Economics of Patents and Standardization: Network Effects, Hold-up, Hold-out, Stacking*, *in* THE CAMBRIDGE HANDBOOK OF TECHNICAL STANDARDIZATION LAW: COMPETITION, ANTITRUST, AND PATENTS 100, 105 (Jorge L. Contreras ed. 2018) (stating that "lack of compatibility . . . is related to network effects," and that "ex post, an implementer is likely to perceive loss of compatibility with their own and others' installed base as a substantial switching cost").

technology, if it switches to a noninfringing alternative, and thus includes sunk costs. If that's all there is to switching costs, however, then one might conclude that the risk of holdup is low whenever those sunk costs themselves are low; but this could lead to an underestimate of the risk of holdup, if the other two potential components are present.

Alternatively, one might view switching costs as distinct from sunk costs, consisting of the *ex post* cost of adopting the noninfringing alternative (the cost of redesign, etc.). As Shapiro and other economists have shown, however—and as the model we introduce in Part III affirms—what actually matters are not the *ex post* redesign costs as such, but rather the difference between the *ex ante* and *ex post* costs of adopting the alternative.[91] Much of the legal (and even some of the economic) literature nevertheless fails to make this point explicitly,[92] and thus risks leading courts and

---

91. *See infra* Part III.A.

92. *See* ANTITRUST ENFORCEMENT, *supra* note 45, at 7:

> If a technology lacks effective substitutes because an SSO chose to include it in a standard, and the costs associated with switching to an alternative standard are high, the owner of patents on that technology may be able to hold up firms wishing to implement the standard by setting higher royalties and less favorable licensing terms than it could have done before the standard was set . . . .

See also FTC REPORT, *supra* note 11, at 5, 189–92, 190 n.53, 222, 225, 225 n.50, 243; Carlton & Shampine, *supra* note 74, at 535 (describing "opportunities for the holders of patents to seek royalty rates that are driven, not by the ex ante incremental value of their technology to prospective licensees compared to other alternatives, but by the incremental market power . . . that inclusion of the technology in the standard confers on the patent holder"); Chao, *supra* note 74, at 303 ("Ex post compatibility value is the value a patent possesses because it allows patent holders to "holdup" a company that has already adopted the patented technology"); Colleen V. Chien, *Holding Up and Holding Out*, 21 MICH. TELECOMM. & TECH. L. REV. 1, 9 (2014) ("[W]hen a patent is asserted after a product is made, the patentee has the upper hand, due not to the economic value of the technology, but instead to the high cost of changing the product to avoid the implicated technology."); Contreras & Gilbert, *supra* note 74, at 1456, 1468, 1488 (stating, *inter alia*, that "[e]x post, a patent owner may be able to charge royalties that reflect the cost of switching to an alternative technology, which may be very high"); Gilbert, *supra* note 74, at 862 ("If the patented technology is adopted after the standard issues and the adopting firm and consumers make investments that are specific to the standard, the cost of switching to the alternative technology is prohibitively expensive"); John M. Golden, *Principles for Patent Remedies*, 88 TEX. L. REV. 505, 518 (2010) ("Strong patent rights might force such a 'trapped' party to pay a licensing fee based more on switching costs than the more intrinsic value

other legal actors to conclude, erroneously, that if the *ex post* cost of redesign is high there is, ipso facto, a substantial risk of holdup. This error could result in an *over*estimate of the risk of holdup in a given case.

On the other hand, much of the existing literature (including even the economic literature) appears to rest on the premise that switching costs can give rise to holdup only if the implementer has incurred substantial asset-specific sunk costs.[93] As we will show, however, if by switching costs one means (increased) redesign costs, this assumption will not always hold true. Rather, sunk costs and increased redesign costs, as well as forgone network benefits, are all *distinct* sources of patent holdup, although each of the three has a common origin in path dependence.[94] To be sure, we suspect that, in most circumstances in which there is a substantial gap between the expected *ex ante* and *ex post* costs and benefits of adopting the alternative, the infringer also will have incurred some sunk costs to implement the infringing technology; such costs would contribute to path dependence to the extent they alter the relative costs and benefits of adopting the alternative.[95]

---

of the patented invention."); Paul J. Heald, *Optimal Remedies for Patent Infringement: A Transactional Model*, 45 HOUSTON L. REV. 1165, 1184, 1187 (2008) ("When the infringer's switching costs are larger than the post-adjudication license fee the patentee can credibly demand, then the infringer will pay to continue to use the patented technology."); Daryl Lim, *Standard Essential Patents, Trolls, and the Smartphone Wars: Triangulating the End Game*, 119 PENN. ST. L. REV. 1, 76 (2014) (stating that "patent owners obtain power over price because of the high switching cost involved."); Michel, *supra* note 74, at 891, 896–97, 904 (2011) (stating, *inter alia*, that "an injunction or its threat can also effectuate patent holdup, leading a manufacturer to pay royalties based on its switching costs rather than the value of the patented technology"); Ted Sichelman, *Purging Patent Law of "Private Law" Remedies*, 92 TEX. L. REV. 517, 522–23, 528, 538, 548, 554–55, 565–67, 571 (2014) (arguing that if the *ex post* cost of redesign is high, there is a substantial risk of holdup).

93. *See* ANTITRUST ENFORCEMENT, *supra* note 45, at 38 n.25 ("The most direct source of switching costs is the difference between the cost of acquiring new infrastructure to implement a new standard and the salvage value of current infrastructure . . ."); FTC REPORT, *supra* note 11, at 5, 22, 58, 79, 144, 222 ("At the time a manufacturer faces an infringement allegation, switching to an alternative technology may be very expensive if it has sunk costs in production using the patented technology."); Stern, *supra* note 70, at 122 (describing a situation in which the holder of a patent withholds the existence of the patent and waits for a producer to sink costs and adopt the technology).

94. *See supra* Part III.A.

95. *See* Durlauf*, supra* note 29 (defining path dependence).

But imagine a situation in which an SSO of which a particular manufacturer is not a member adopts a standard, and the manufacturer now wants to make and sell standard-compliant devices. The manufacturer may not have incurred any sunk costs prior to standard adoption, but *ex post* the costs and benefits of adopting an alternative technology are much different from what they would have been *ex ante* (prior to standard adoption), and the manufacturer may be subject to holdup as a result.[96] As in the first example, however, we worry that in such a case courts or other legal actors might underestimate the risk of holdup, in this instance on the basis of a misperception that high sunk costs are a necessary condition for holdup to arise.

III. Modeling Patent Holdup

In Section A of this Part, we present a simple model that identifies and isolates the three sources of path-dependent patent holdup. In addition, we will show that: (1) patent holdup can arise from any of the three sources, and thus that nonsalvageable sunk costs are not a necessary condition; (2) high switching costs, as such, are irrelevant; and (3) what is relevant is the difference between the expected *ex ante* and *ex post* costs and benefits of adopting the alternative—or, to put it another way, the infringer's opportunity cost of *not* having chosen the noninfringing alternative *ex ante*. As a result, conventional understandings of patent holdup that gloss over these points pose of a risk of either over- or underestimating holdup risk under a given set of circumstances. In Section B, we present some illustrative examples of when a risk of holdup may, or may not, be present.

A. The Model

Assume that there are two (substitute) technologies, indexed by $t = 1, 2$, and that there are two stages, *ex ante* and *ex post*. In

---

96. Another possible example would be one in which all or most of the sunk costs are salvageable, that is, can be easily redeployed to make use the noninfringing alternative rather than the patented technology.

the *ex ante* stage, an implementer can choose to adopt one of the technologies. Adoption of technology t requires an upfront fixed cost investment of $K_t = \alpha + k_t$, where $\alpha, k_t \geq 0$. Here k_2 gives the technology-specific component of K_t, while α describes any joint cost savings that could be redeployed to the other technology. Choosing technology t *ex ante* will provide a variable profit (total revenue minus total variable cost) of π_t,[97] providing a total profit of $\pi_t - K_t$, provided that the implementer does not switch technologies *ex post*.

Let t^* denote the technology adopted *ex ante*, and let $t = 3 - t^*$ denote the alternative. The *ex post* stage occurs just after t^* has been adopted, and surrounds potential switching to t. The joint cost α can be redeployed to t, and thus need not be reincurred. Instead, the implementer must pay only a t-specific fixed cost $k_{t|t^*}$ —the *ex post* adoption cost of t *conditional on ex ante adoption of* t^*. Similarly, conditional on *ex ante* adoption of t^*, the implementer will obtain variable profit of $\pi_{t|t^*}$ if it switches to t *ex post*. The notation reflects that we are allowing the costs and revenues of adopting a given technology to depend on whether it was adopted *ex ante*, or switched-to *ex post*. We assume *ex ante* expectations are accurate, so that $\pi_{t^*|t^*} = \pi_{t^*}$ and $k_{t^*|t^*} = 0$, where the latter reflects that the fixed cost of adopting t^* need not be incurred again if the implementer declines to switch.

We say that *ex ante* commitment to t^* generates *path dependence* if either $k_{t|t^*} > k_t$, $\pi_{t|t^*} < \pi_t$, or $k_{t^*} > k_{t^*|t^*} = 0$. Such conditions, if satisfied, reflect some dynamic change in the relative commercial value of t that is causally attributable to *ex ante* adoption of t^*. This can happen if there is a reduction in the overall profitability of the alternative ($k_{t|t^*} \neq k_t$ or $\pi_{t|t^*} < \pi_t$), and these effects can be regarded as opportunity costs of not having adopted the alternative *ex ante*. Alternatively, in the more familiar case, the change in relative value may arise because sunk costs have left the patented invention with higher marginal value ($k_{t^*} > 0$). We provide some intuitive examples in Section B below.

To model patent holdup, suppose that technology 1 is patented while technology 2 is not,[98] and that technology 1 is superior in

---

97. This can be interpreted as an expected present discounted value.

98. If both technologies are patented, there is price competition between the patented technologies (both *ex ante* and *ex post*). But, to the extent this results in different fee offerings for the alternative *ex post* versus *ex ante*, this is attributable

that it provides a greater ante total profit ($\pi_1 - K_1 > \pi_2 - K_2$). In the model presented here, we suppress patent uncertainty and simply assume that litigation expectations are sufficient to ensure that the implementer would rather take up the alternative than infringe and litigate.[99]

If the implementer and patentee bargained *ex ante*, the maximal fee the patentee could charge would be

$$\mathcal{F}^A = \pi_1 - k_1 - [\pi_2 - k_2]$$

Thus, the maximal *ex ante* fee is increasing (decreasing) in π_1 and k_2 (k_2 and π_2) and invariant in α. Holdup can occur when licensing negotiations do not occur until the *ex post* stage. The fixed cost k_1 is now sunk and, at the margin, the costs or benefits of technology 2 may have changed as a result of the prior commitment to 1. Conditional on this *ex ante* choice, the maximal fee the patentee can obtain through *ex post* negotiations is

$$\mathcal{F}^P = \pi_1 - [\pi_{2|1} - k_{2|1}]$$

We define the *holdup rents* as the difference between $\mathcal{F}^A$ and $\mathcal{F}^P$, or the increase in the implementer's willingness to pay for a license when licensing terms are not negotiated until *ex post*.

---

to the same factors that facilitate holdup, as identified below. When the two *ex ante* technologies are patented, there is at present no theoretical consensus on what the appropriate royalty should be. *See* Norman V. Siebrasse, *Holdup, Holdout, and Royalty Stacking: A Review of the Literature*, *in* COMPLEX PRODUCTS, *supra* note 6, at 239, 241–42.

99. In the Technical Appendix, we endogenize patent uncertainty (the patentee's winning probability) and remedies. Our findings about holdup and path dependence do not change, and the changes that do arise are already explicated in Lemley & Shapiro. Lemley & Shapiro *supra* note 11, at 2035–39. The biggest change, as Lemley and Shapiro explain, is that the probabilistic nature of patent validity and infringement can result in license fees being systematically excessive, which they refer to as an "overcharge," even when negotiated *ex ante*. *See id.* at 2002. This overcharge is separate from holdup, which, as we state above, refers to an increase in the *ex post* royalty rate as compared with the rate that could have been negotiated *ex ante*, that does not result from an increase in the anticipated value of the patented technology. Because the overcharge identified by Lemley and Shapiro also arises *ex ante*, it does not represent holdup in that sense.

Proposition. *Holdup rents are equal to*

$$\mathcal{H} = k_1 + \Delta\pi_2 + \Delta k_2$$

where $\Delta\pi_2 \equiv \pi_2 - \pi_{2|1}$ *and* $\Delta k_2 \equiv k_{2|1} - k_2$.

This simple equation clarifies what phenomena ultimately contribute to patent holdup. Consistent with classic (nonpatent) holdup, the first term is simply the technology-specific sunk cost investment k_1 in the patented technology. This reflects a form of path dependence, because the choice to invest in k_1 alters the relative value of the technologies going forward. The terms $\Delta\pi_2$ and Δk_2 reflect potential opportunity costs of adopting technology 2 *ex post* rather than *ex ante*. This demonstrates that the relationship between technology-switching and holdup is ultimately driven by path dependence effects.

Corollary 1. *Holdup rents are positive if and only if* ex ante *commitment to the patented invention generates path dependence.*

Corollary 2. *The switching cost* $k_{2|1}$ *does not independently reveal anything about the magnitude of the holdup rents.*

Hence, even if the redesign cost $k_{2|1}$ is quite large, it influences holdup only to the extent that it exceeds k_2, the cost of adopting technology 2 *ex ante*. Thus, holdup rents are not increasing in the size of the redesign cost, but rather in the size of the opportunity cost Δk_2, which captures the increase in the cost of adopting the alternative. On the other hand, even if redesign costs and sunk costs are zero, holdup may still arise, for *ex ante* commitment to technology 1 may diminish the post-adoption variable profitability of technology 2. That is, $\pi_{2|1}$ may fall below π_2. An immediate implication is that it is a mistake to regard large fixed costs as an essential ingredient of patent holdup.[100]

---

100. Apart from litigation uncertainty and bargaining strength, which we ignore, the holdup royalty identified by Shapiro, s*upra* note 11, at 287–303, turns on the cost of redesigning the product to avoid using the patented feature. Shapiro denotes this redesign cost as "F," and it is clearly a forward looking redesign cost,

As one would expect, in the absence of path dependence effects distorting the value of the alternative (i.e. when $\Delta\pi_2 = \Delta k_2 = 0$), our model reduces to the standard point that *ex post* opportunism allows the patentee to extract the implementer's technology-specific sunk cost investment.

B. Examples

1. Sunk Costs

For an example of patent holdup attributable to sunk costs alone, consider the facts of *Riles v. Shell Exploration & Production Co.*[101] Riles owned a patent on a method for installing the foundation for an oil drilling platform, which Shell was found to have infringed.[102] Shell's evidence suggested that the value of the invention over alternative nonpatented methods was about $350,000.[103] The *ex ante* royalty ($\mathcal{F}^A$) therefore would have been no more than this amount. Shell infringed the patent in laying the foundation for a platform, and then built a platform atop the foundation at a cost of $84 million.[104]

In the actual case, it appears that Riles did not seek an injunction, but did argue that he was entitled to a reasonable royalty based on what Shell would have been willing to pay if the district court *had* enjoined Shell's use of the platform.[105] That

---

id. at 287, corresponding to the cost we denote by $k_{2/1}$—the cost of switching to the alternative given that the patented technology was adopted. In the case Shapiro is concerned with, when the redesign cost for the alternative is higher than it would have been *ex ante*, and differential profits from the alternative technology are zero, then $k_{2/1}$ can be expressed as $k_2 + D$. That is, in Shapiro's model, holdup turns on $F = k_2 + D$. In our model, in contrast, in this scenario the holdup rent $H = k_1 + D$. The reason for this apparent difference is that Shapiro implicitly assumes that $k_1 = k_2$, in which case the two expressions are equal.

101. 298 F.3d 1302 (Fed. Cir. 2002).

102. *Id.* at 1305–06.

103. *Id*. at 1313.

104. *Id*. at 1311.

105. *See id.* (stating Riles' expert's theory: "Shell's construction of the platform with a patented method could result in an injunction . . . . If such an injunction were in place . . . Shell would have had to choose between abandoning

amount would have dwarfed the $350,000 *ex ante* value of the invention: assuming, for lack of any evidence one way or the other, that the costs and benefits of adopting the nonpatented method did not change,[106] *ex post* Shell would be better off paying a royalty up to just under $84,350,000, rather than abandoning the project altogether. In terms of Proposition 1 above, $k_1 = \$84,000,000$, $\Delta\pi_2 = 0$, and $\Delta k_2 = 0$, so $\mathcal{H} = k_1 + \Delta\pi_2 + \Delta k_2 = \$84,000,000$. Fortunately for Shell, the court rejected the argument that Riles could have enjoined Shell's use of the platform, notwithstanding Shell's use of the patented method to construct the platform, and concluded that the royalty the jury actually awarded (a mere $8.7 million) was excessive.[107]

Riles thus illustrates how, in a case involving large, nonsalvageable sunk costs, an injunction could enable a patent owner to extract an excessive royalty *ex post*, as well as how courts can employ this understanding to rein in excessive damages awards. Of course, each case must turn on its own facts, and sometimes it may not be as easy as it was in *Riles* to determine what those sunk costs are. In cases involving technical standards, for example, the evidence might show that most of the cost the infringer incurs to implement a particular feature could be redeployed if it were to implement a noninfringing alternative; or, conversely, it might show that not only the cost of deploying the feature, but also a range of complementary investments in other standard-compliant features would be lost if the infringer switched or abandoned.[108] Nevertheless, it may not always be necessary to determine the value of the sunk costs with precision. If the *ex ante* value of the technology can be estimated with tolerable accuracy,

---

its $84 million platform or paying Riles a percentage").

106.   If Shell had to demolish the existing platform in order to install a new one using the nonpatented method for laying the foundation, the *ex post* redesign costs almost surely would exceed *ex ante* costs. *See supra* Part II.B. Our example above abstracts away from this contingency.

107.   *See id*. at 1311–13 (rejecting Riles' argument and declaring the jury award excessive).

108.   The discussion of sunk costs should note that sunk costs may refer to the investment in the industry as a whole, not just technology to the particular standard, when it comes to a standard. *See, e.g.*, Carlton & Shampine, *supra* note 74, at 542 (describing a situation in which collective voting can allow a patent holder to hold up an individual firm within an industry); Farrell et al., *supra* note 21, at 616 (noting that holdup can be especially severe for industry standards).

and if (by any measure) the sunk costs are disproportionate to that value, then it is reasonable to conclude that a serious holdup risk is present.[109]

2. Increase in Costs of Adopting the Alternative

As noted above, commentaries on patent holdup frequently highlight the costs of switching to an infringing product to replace the patented component with a noninfringing alternative. What actually matters, though, is not the cost of adopting the alternative, as such, but rather the amount by which this cost post exceeds the cost of adopting the alternative *ex ante*—in terms of Proposition 1 above, $\Delta k_2 (\equiv k_{2|1} - k_2)$, not $k_{2|1}$ itself.

To illustrate, imagine that the patented and alternative technologies are plumbing valve configurations.[110] *Ex ante*, a contractor must decide which valve type to install, and it will make this decision at an intermediate stage of the home's construction, before the walls and flooring are installed.[111] This installation cost is k_t for valve technology t. For simplicity, the valves are incompatible but equivalent; one valve works as well as the other, so the value of the house is the same regardless of the choice, and the valves themselves cost the same, so the choice is arbitrary.[112] If the contractor had to switch *ex post*, after the home's construction was complete, it would have to temporarily tear down the walls and floors to get at the plumbing; this deconstruction imposes its own cost $D > 0$, which is independent of the valve technology.[113] Thus, if technology 1 is adopted *ex ante* and its use

---

109. This is not necessarily to say that an injunction should be refused, as that would require consideration of other factors, such as whether *ex ante* bargaining would have been feasible, and how difficult it is to accurately assess a reasonable royalty in lieu of an injunction. *See* Siebrasse et al., *supra* note 7, at 144–59.

110. This example is adapted from a famous case dealing with contract remedies and their potential to facilitate holdup. *See* Jacob & Youngs, Inc. v. Kent, 230 N.Y. 239 (1921).

111. *Id.* at 240.

112. *Id.*

113. *Id.*

were to be enjoined *ex post*, the *ex post* cost of switching to technology 2 would be $k_{2|1} = k_2 + D$, and the holdup royalty $\mathcal{H} = k_1 + D$, or the sunk cost of installing the original valve plus the heightened cost of redesign.[114]

Note, however, that redesign contributes to holdup only to the extent that it comprises some *extra* cost that would not have arisen if the implementer had originally designed its product around the alternative technology. The cost of implementing the alternative technology, k_2, does not represent holdup, even though it is a cost that has to be incurred to switch, simply because that cost would have been incurred even if technology 2 had been adopted in the first place.

To see why this is important, modify the hypothetical to assume that the two valves still work equally well, but valve technology 1 is easier and cheaper to install, so that $k_1 < k_2$. In that case, the forward-looking "switching costs" (in the intuitive sense of the cost of switching to the new technology), or $k_2 + D$, will be very high, because of the costs of installing valves with technology 2. But the holdup royalty is still $\mathcal{H} = k_1 + D$. The difference between the two, $k_2 - k_1$, represents the value of technology 1 over technology 2, and there is nothing wrong with allowing the patentee to extract some part of that difference. Thus, the fact that the forward-looking cost of switching is high does not in itself signal patent holdup. It is only D, the *increased* cost of implementing technology 2, which contributes to holdup, along with the sunk cost of having adopted technology 1.

---

114. Perhaps more realistically, one might also expect there to be some nonsalvageable sunk costs, the recovery of which would be forfeited if the defendant had to demolish portions of the building to access the plumbing, and that the holdup royalty would reflect some portion of these costs as well. *Cf. supra* note 106 (noting that our *Riles* example abstracts away from potential redesign costs). The broader point is that, while intuition supports the notion that sunk costs are past costs and redesign costs are future costs, there may be no sharp dividing line between what is a nonsalvageable sunk cost and what is a redesign cost, when the process of redesign entails demolishing something that already has been constructed.

3. Forgone Benefits

The third component of patent holdup is perhaps most likely to arise in cases involving SEPs or other patents exhibiting strong network effects. To see how this third component by itself can contribute to patent holdup, consider the following stylized example. (We provide a more rigorous example in the Technical Appendix.) Suppose that, *ex ante*, an SSO is choosing between two standards, one of which incorporates (among other features) a software-patented technology and the other a nonpatented technology. The SSO adopts the first standard, even though the patented technology at issue is no better than the nonpatented. (In other words, $\pi_1 = \pi_2$.) For simplicity, suppose further that either technology can be implemented by means of a costless firmware update, so that $k_1 = k_2 = 0$. The *ex ante* royalty for the patented technology therefore would be zero: $\mathcal{F}^A = \pi_1 - k_1 - [\pi_2 - k_2] = \pi_1 - \pi_2 = 0$.

A non-SSO member implementer then inadvertently infringes. Suppose further that the *ex post* cost of adopting the nonpatented technology remains zero (that is, $k_{2|1} = k_2 = 0$), but that switching to that alternative *ex post* means that the implementer will earn zero profit on devices that incorporate the nonpatented technology, because those products will no longer interoperate with other producers' (that is, $\pi_{2|1} = 0$). The *ex post* royalty therefore reduces to: $\mathcal{F}^P = \pi_1 - \pi_{2|1} = \pi_1$, and the holdup portion to $\mathcal{H} = k_1 + \Delta\pi_2 + \Delta k_2 = \pi_2$.

Counterintuitively, then, what contributes to holdup in this example is not the variable profit that is lost from not selling the patented technology (π_1), but rather the diminution in the variable profit attributable to the *ex post* use of the noninfringing alternative (π_2)—that is, the forgone benefit attributable to the SSO *not* having chosen the nonpatented alternative *ex ante*.

To be sure, in the above example we assumed that $\pi_1 = \pi_2$, so the variable profit that the implementer loses from not selling the patented technology (π_1) is the same amount as the diminution in the variable profit attributable to the *ex post* use of the noninfringing alternative (π_2). But this will not always be the case. Suppose instead that the SSO chose the first standard because the patented technology really is superior to the nonpatented

alternative, that is, that $\mathcal{F}^A = \pi_1 - k_1 - [\pi_2 - k_2] = \pi_1 - \pi_2 > 0$. The *ex post* royalty will still be $\mathcal{F}^P = \pi_1 - \pi_{2|1} = \pi_1$, and the holdup portion will still be $\mathcal{H} = k_1 + \Delta\pi_2 + \Delta k_2 = 0 + \pi_2 + 0 = \pi_2$. In this case, however, only part of the *ex post* royalty is attributable to holdup; the other part reflects the *ex ante* value of the technology $(\pi_1 - \pi_2)$.

4. Injunctions and "Lag Time"

Another variation on this theme involves "lag time" and injunctive relief. In their paper *Patent Holdup and Royalty Stacking*, Mark Lemley and Carl Shapiro provide a model in which holdup can arise because an injunction, if enforced, would result in a period of time during which the implementer is unable to sell anything, pending the development of a noninfringing product.[115] This is very similar to the case of redesign costs, but rather than being a fixed cost, the lag time results in forgone variable profits.[116]

To see why, we can normalize a product's lifespan to 1, and let λ denote the length of the lag period $(0 < \lambda < 1)$, expressed as a portion of the product lifetime. Then, as a result of *ex ante* adoption of technology 1, the lifetime profits from a product utilizing technology 2 fall from π_2 to $\pi_{2|1} = (1 - \lambda)\pi_2$. Ignoring fixed costs, lag time in technology-switching generates holdup rents of $\mathcal{H} = \lambda\pi_2$. Intuitively, although most patented technologies contribute only incrementally to a final product, an injunction is usually not an incremental remedy: it typically requires the defendant to withdraw the entire product for some amount of time, even if only a small component is infringing.[117]

What unifies all these examples is that the holdup comprises an opportunity cost of not having chosen the alternative technology *ex ante*.[118] The differential profit in this example parallels the

---

115.   *See* Lemley & Shapiro, *supra* note 11, at 1997, 2038 (providing an example of holdup due to an injunction).

116.   *Id.*

117.   *See supra* Part II.A.

118.   Indeed, one could readily consider the (negative) *ex post* benefit of switching to an alternative as a (positive) *ex post* switching cost, though we think it is somewhat more intuitive to distinguish the redesign costs from forgone benefits, as above. *See supra* note 20 and accompanying text.

differential cost in the example of the plumbing valves. What matters is not the cost (or diminished profit) of adopting technology 2, but the difference between that forward-looking cost (or profit) and the cost (or profit) that would have been incurred had technology 2 been adopted in the first place. When the two technologies are roughly equivalent, the difference will be small, and most of the forward-looking penalty from switching will indeed contribute to holdup. But when technology 1 is superior, some substantial part of the forward-looking costs, or lost profits, will represent the true value of the technology, not holdup. And forward-looking costs are not necessary to holdup at all. The sunk costs of adopting technology 1 that contribute to holdup are also a consequence of not having chosen technology 2 in the first place, which would have avoided those sunk costs, and consequently they contribute to holdup in any event.

5. More Complex Cases

The foregoing examples were intended to isolate, as much as possible, the three different sources of holdup we have identified, but real-world fact patterns will more typically present two or even all three of the potential sources of patent holdup, at least to some degree. Consider, for example, a typical case brought by a patent assertion entity (PAE). Most commonly, PAEs assert software or business method related patents that cover technologies incorporated into complex devices that integrate many complementary features.[119] This fact alone can make the process of redesigning the device *ex post* more complicated than it would have been *ex ante*, insofar as a switch would entail not only losing one's sunk investment in the infringing technology, but also (as in the plumbing example) incurring the cost of uninstalling much of what already has been done. In addition, PAEs rarely transfer

---

119.   *See* FTC STUDY, *supra* note 2, at 5, 74–75, 135 ("Of all the patents held by PAEs in the FTC's study, 88% fell under the Computers & Communications or Other Electrical & Electronic technology categories, and more than 75% of the Study PAEs' overall holdings were software-related patents."); Brian J. Love et al., *An Empirical Look at the "Brokered" Market for Patents*, 83 MO. L. REV. 359, 399–406 (2018) ("[L]ooking closer at the technology categories for which we did see a large number of packages, software stands out as the most interesting.").

technical know-how to their targets, preferring instead to target firms that have independently invented the allegedly infringing technology,[120] thus creating a risk of *ex ante* opportunism. On the other hand, PAE cases that are litigated to judgment in the United States rarely result in the entry of an injunction,[121] and so the risk of enabling the owner to extract a royalty reflecting the differential cost of adopting the alternative is mitigated. Courts in the U.S. would still need to ensure, however, that the royalties awarded do not reflect this cost or, as above, the infringer's sunk cost.

SEP cases also can present enormous differential *ex post* redesign costs, unlike our stylized example in the preceding subsection. If a firm is found to have infringed an SEP and is subject to an injunction, it faces a choice among (1) abandoning the project entirely (and thus forfeiting its sunk costs, as above); (2) switching to a noninfringing alternative; or (3) paying what the SEP owner demands. Often, switching will entail having to redesign many complementary features, even if there is some alternative standard available to switch to—and, as discussed in the preceding subsection, the benefits to be derived from employing that alternative also are likely to be much lower *ex post* than they would have been had the alternate been chosen (by the infringer and others) *ex ante*. Another possibility is that the infringer could convince all or most of the other firms that implement the patented technology to switch to an alternative standard, but the *ex post* cost of such an en masse switch would surely exceed the *ex ante* cost of adopting the alternative in the first place.[122]

---

120.   *See* COTTER, *supra* note 1, at 211–12, 215–16; Love et al., *supra* note 119, at 404–05 (finding that the brokered patent market is a market for the transfer of potential legal liability, not for the transfer of technology).

121.   *See* Christopher B. Seaman, *Permanent Injunctions in Patent Litigation After* eBay: *An Empirical Study*, 101 IOWA L. REV. 1949 (2016) (noting that PAE cases brought in the United States rarely result in an injunction). Prior to the U.S. Supreme Court's decision in *eBay Inc. v. MercExchange, L.L.C.*, 547 U.S. 388 (2006), courts almost always granted the prevailing patent owner an injunction. The pre-*eBay* standard of near-automatic injunctive relief largely remains the rule outside the United States, except in SEP cases. For discussion, see, for example, Contreras et al., *supra* note 6, at 175–84 (discussing the law regarding injunctions in the European Union); Siebrasse et al., *Injunctive Relief*, *supra* note 7, at 134–41 (discussing the law regarding injunctions in England).

122.   *See* Carlton & Shampine, *supra* note 74, at 542 (presenting a situation in which the *ex post* cost exceeds the *ex ante* costs of adopting an alternative).

Similarly, a phenomenon known as "patent ambush" provides another example of how sunk costs, as well as potentially the other two costs, can contribute to holdup.[123] Patent ambush can arise when, for example, an SSO member induces other members into adopting a standard that (unbeknownst to the other members) is or will be infringed by the member's patents, and then once those other members are locked in demands an *ex post* royalty that exceeds the patent's *ex ante* value.[124] Revising the standard *ex post* would result in a waste of, at the least, much of the time and effort that went into developing the standard *ex ante*, and thus we would expect k_1 to be substantial.[125] In addition, however, since the technology is standardized, it is likely that the *ex post* costs of shifting to an alternative also would be higher than the *ex ante* costs of doing so, and that the *ex post* benefits any single firm would derive from unilaterally switching would be lower than the *ex ante* benefits had an alternative standard been chosen.[126]

6. Further Observations

The foregoing analysis identifies the kinds of costs that are relevant to holdup. While the problem of sunk costs has long been well understood, we have argued that switching costs, in the intuitive sense of what it would cost to switch to the alternative technology, are not relevant in themselves. Instead, what is relevant is the differential switching cost—how much more would it have cost to switch than it would have cost to implement the alternative. Similarly, the question is not how much less profit the infringer would make if it had to switch to an alternative, but how much less it would make compared with what it would have made had it adopted the alternative initially. While our analysis has

---

123. Cotter, *supra* note 11, at 1188–89.

124. *Id.*

125. Most observers agree that when such conduct occurs the deceived members are entitled to some form of legal redress, though opinions differ as to whether the optimal legal response lies in antitrust, contract, or—by denying injunctive relief, or possibly withholding any form of relief under the doctrine of equitable estoppel—patent law itself. *Id.* at 1188–1200.

126. *Id.*

aimed to clear up some conceptual fuzziness, it also has practical implications. For example, many authors have suggested that the availability of injunctive relief should be conditioned, at least in part, on the potential for holdup if the injunction is granted.[127] A prior question, then, is whether there is a potential for holdup on the facts of the case. Our analysis implies that determining the holdup risk does not always amount to simply asking whether the cost of switching would be high, as a naïve understanding of "switching costs" might imply. Instead, it is the differential costs which are relevant.

To be sure, there may be practical problems in implementing this analysis, assuming it is persuasive. Our analysis does imply that more information is needed to assess holdup than under a model which simply looks at forward-looking switching costs, for example; and it might be difficult for courts to determine the values of the relevant variables in some cases. But if our analysis is right, it would not make sense for a court to assess holdup by taking into account only forward-looking switching costs, simply because it does not have the evidence to determine the differential switching costs. The better approach would be for the court to recognize that it does not actually know the potential for holdup when deciding whether an injunction should be granted. Further, it may not always be so difficult to obtain the information needed to determine whether a holdup risk is present: all that is necessary is some basis for concluding that the *ex post* cost of adopting the alternative is substantially greater than the *ex ante* cost, without necessarily quantifying either.

Similarly, in the redesign with lag time case presented above, all that would be necessary is some way of ascertaining how long an injunction would result in the defendant being excluded from the market, and some rough estimate of the *ex ante* profit the defendant would have expected to earn had it employed the alternative initially. The greater these two variables are, the greater the risk of holdup. On the other hand, to the extent that a court would insist on quantitative evidence of the holdup costs, the additional evidence required by our approach might be easy to

---

127.   *See supra* note 11 and accompanying text (gathering sources).

obtain. If a party were able to submit specific evidence of how much it would cost to implement an alternative technology, it should not be too difficult also to provide evidence as to how much it would have cost to have implemented it initially. Moreover, courts can only deal with the evidence before them, and our analysis gives the parties an incentive to produce the right kind of evidence. If an infringer argues in its pleadings that an injunction should be denied because of high costs to switch to the alternative, our analysis suggests that the patentee should bring forth evidence that the costs would have been just as high *ex ante* (assuming, of course, that that is true).

Finally, we note that, although redesign costs as such are irrelevant to holdup, it is possible that in some cases they would be of approximately the same magnitude as the cost already sunk in the patented technology. In such a case, redesign costs could serve as a proxy for sunk costs, if the latter are (for reasons suggested in the preceding section) difficult to determine but it is reasonable to believe that the sunk cost of adopting either technology would be roughly equal. Of course, whether redesign and sunk costs actually *are* roughly equal is an empirical matter that likely would vary from case to case.

IV. Some Possible Critiques

Although many observers have found the economic analysis of patent holdup as developed by Shapiro and others persuasive, a distinct minority of observers (including the new head of the DOJ's Antitrust Division) has expressed serious reservations over both the theory and its relevance in addressing real-world problems.[128] The critiques generally fall into three categories: first, that the theory presents an illogical extension of classic holdup theory; second, that patent holdup cannot be a real-world phenomenon, because the industries that one might expect to have been most affected by it (e.g., smartphones) are among the most innovative in

---

128. *See supra* note 9 and accompanying text (citing speeches made by Makan Delrahim, Assistant Attorney General, Antitrust Division, in which he critiqued courts' reluctance to grant injunctions in SEP cases).

all of human history; and third, that a practice known as "reverse holdup" or "holdout," which occurs when implementers negotiate in bad faith for the right to use patented technologies, is a form of opportunism that merits greater concern than patent holdup. To the extent some or all of these claims are based on empirical assertions about the frequency with which holdup occurs (or not) in the real world, our analysis neither proves nor disproves them. We think it does show, however, that there is nothing logically unsound about the patent holdup concept. Moreover, because any of the three components of patent holdup can exist independently of the others, the conditions that enable holdup can arise in a number of commonplace settings; holdup, in other words, does not have to devastate an entire industry to be a matter of appropriate concern. In addition, we will show that there is nothing unusual about taking holdup considerations into account in crafting intellectual property (IP) rules and standards, but rather that several other doctrines in patent and other IP law can be viewed as means for preventing the opportunistic exploitation of IP rights.

A. Is Patent Holdup Really Holdup?

As our discussion throughout has indicated, other commentators have explored a wide variety of circumstances in which holdup can occur. The analysis presented in Parts II and III shows that the general principle linking all of these examples is the path dependence that arises when an implementer changes its position after having adopted a patented technology without first having come to terms with the patent owners. The resulting holdup rent that the patent owner can extract can then be broken down into the sum of the implementer's sunk costs plus the opportunity cost of not having chosen the next-best nonpatented technology *ex ante*. As we also noted above, however, this type of situation is not identical to the descriptions of holdup found in the classic holdup literature, which typically models one party to an incomplete contract extracting *ex post* concessions from the other—concessions that are unanticipated *ex ante*, due to the party's "guile" in concealing its true intentions.[129] Patent holdup

---

129.	*See supra* notes 75–76 and accompanying text.

skeptics sometimes argue that, because patent holdup does not fit the classic pattern, there is something wrong with patent holdup theory.[130]

We disagree. Both the classic and the patent holdup literature demonstrate how a party that finds itself locked in, based on some initial, not easily reversible choice, can become vulnerable to another's ability to leverage the extraction of quasi-rents *ex post*.[131] Moreover, while some analysts emphasize guile as a necessary condition for classic holdup, not all of them do; and in the settings of relevance to patent holdup, the fact that implementers sometimes can anticipate the risk of being held up hardly eliminates the harm to static and dynamic efficiency if and when patent holders extract royalties that exceed the value of their inventions.[132] Indeed, even Galetovic and Haber appear, perhaps inadvertently, to concede the point.[133] As noted above, in their discussion of patent holdup's alleged inconsistencies with classic holdup, they argue that implementers know they bear the risk of inadvertently trespassing on patents, and thus can "either insure themselves against that risk (by setting aside a reserve) or demand a higher expected return on capital, as with any other business risk."[134] Either response, however, increases the cost of investing in a holdup-prone industry, and thus inefficiently diverts resources to other areas with a lower social rate of return but less holdup risk.[135] This consequence is precisely why patent holdup reduces social welfare.[136]

Finally, while Galetovic and Haber argue that, in an otherwise perfectly competitive market, one party's extraction of quasi-rents cannot be a long-term equilibrium because the other party will be unable to cover its fixed costs and thus eventually will go out of business,[137] that theoretical observation hardly eliminates the

---

130. *See, e.g.*, Galetovic & Haber, *supra* note 78, at 13–28.

131. *See, e.g.,* Klein et. al, *supra* note 13, at 299–300 (discussing opportunistic leverage of quasi-rents).

132. *See supra* note 78 and accompanying text.

133. Galetovic & Haber, *supra* note 78, at 27.

134. *Id.*

135. *Id.*

136. *See supra* note 13 and accompanying text.

137. Galetovic & Haber, *supra* note 78, at 28.

possibility that holdup can occur, in the short run, in real-world, imperfectly competitive markets.[138] We doubt that Shell would have gone out of business, for example, even if Riles had been able to extract an unwarranted $84 million royalty on one occasion, but that does not mean that courts should turn a blind eye to the potential for such abuses in real cases. More generally, one cannot credibly claim to have "disproven" the validity of patent holdup as a general concept by imposing strong assumptions on market structure (namely that it is perfectly competitive) serving effectively to guarantee that holdup would not be a viable means of extracting revenues.

B. Patent Holdup Is Not a Problem, Because It Is Not Systemic

A second, related argument is that there is no empirical evidence of patent owners engaging in pervasive, systemic patent holdup in the very industries holdup theorists are most concerned with (e.g., telecommunications).[139] Indeed, according to the critics, if holdup were pervasive one would expect innovation and growth in the affected industries to "stagnate, wither, or die,"[140] whereas if one looks "across human history, it is not clear that the commercialization of complex technologies has ever been faster than it is today in those industries that reform proponents point to as most plagued by the patent holdup 'problem.'"[141]

---

138. *Id.*

139. For citations to sources making this claim, see Jorge L. Contreras, *Much Ado About Hold-Up*, 2019 U. ILL. L. REV. 875, 878 n.10 (2019); J. Gregory Sidak, *The Antitrust Division's Devaluation of Standard-Essential Patents*, 104 GEO. L.J. ONLINE 48, 61 n.49 (2015); Putnam, *supra* note 74, at 1004.

140. Galetovic & Haber, *supra* note 78, at 28.

141. Alexander Galetovic et al., *Patent Holdup: Do Patent Holders Holdup Innovation?*, HOOVER IP² WORKING PAPER SERIES, May 2014, at 1, 6. *See also* Alexander Galetovic et al., *An Empirical Examination of Patent Holdup*, 11 J. COMPETITION L. & ECON. 549, 572 (2015) [hereinafter *Empirical Examination*] (concluding that "the rate of innovation—as reflected in quality adjusted relative prices—has rarely, if ever, been faster than it is today in exactly those products that scholars agree are theoretically subject to SEP holdup"); David J. Kappos, *The *Real* Innovation Economy—Debunking Myths at the Intersection Between Intellectual Property and Competition Laws*, https://perma.cc/C6HT-UL8B (PDF) ("Mythtellers would have us believe that patent owners are wielding SEPs offensively to hold up standards implementers . . . . But if this were the case, the cost of popular consumer technology would not be plummeting while new

Although we agree that whether, or to what extent, patent holdup occurs in the real world is ultimately an empirical matter, the implication that patent holdup is a problem only if it is "pervasive" or "systemic" is a non sequitur.[142] If our analysis above is correct—that the ability to engage in patent holdup depends on path dependence, that settings conducive to patent holdup are not uncommon, and that the three components of a holdup royalty can exist independently of one another—patent holdup does not have to be systemic to be capable of reducing social welfare. Seeing how the empirical critiques of patent holdup do "not claim[ ] that individual firms never attempt to engage in behavior that can be characterized as holdup,"[143] the conclusion that holdup is not systemic may well be accurate, for all we know, while still being of any limited relevance for purposes of determining whether injunctive relief should issue on the facts of any one particular case.[144] If the choice were between always granting an injunction without tailoring or conditions, and never granting any form of injunctive relief, perhaps the question of whether holdup was systemic, at least in a particular industry, would be central. But the traditional approach to injunctive relief looks to the facts of the particular case.[145]

---

features simultaneously spring up with equal fervor."); Putnam, *supra* note 74, at 967:

> Despite having launched arguably the most successful consumer product in history, Apple even claims to have 'faced excessive royalty demands, onerous contract terms and the threat of injunctions barring the sale of a revolutionary new product,' 'a history . . . [that] has left [the FRAND licensing] promise at least partially unfulfilled.' . . . Faced with these economic headwinds, Apple has sold just 1.2 billion iPhones in 10 years, worth $738 billion. (citations omitted).

142. *See* Galetovic et al., *Empirical Examination*, *supra* note 141, at 555 ("[W]e are not claiming that individual firms never attempt to engage in behavior that can be characterized as holdup. Rather, we do not find evidence consistent with the hypothesis that SEP holdup is a systemic problem.").

143. *Id.*

144. *Id.*

145. *See* eBay Inc. v. MercExchange, L.L.C., 547 U.S. 388, 391 (2006) (providing that a plaintiff seeking injunctive relief must demonstrate irreparable injury; that other damages available at law, such as monetary damages, are inadequate; that a remedy in equity is warranted; and the public interest would not be disserved by a permanent injunction).

Further, rather than the absence of patent holdup serving as a reason for courts to enter injunctions in SEP, PAE, and other cases, it may be that case law imposing limits on the entry of injunctions is itself a leading factor constraining firms from engaging in holdup.[146] Again, the question ultimately is an empirical one, but for now we cannot rule out the possibility that legal reforms were necessary to prevent patent holdup from getting worse.

C. Holdout Is Worse Than Holdup

A third argument is that, because the patent owner has incurred *ex ante* sunk costs in the research and development that resulted in a patented invention, the owner is subject to a form of *ex post* opportunism ("holdout") on the part of implementers who refuse to pay fair compensation; and that this opportunism is, if anything, more insidious than patent holdup, because absent a fair reward the incentive to engage in innovation withers away.[147] Implicit in this argument are two related points that need to be unpackaged: first, that holdout is itself a form of holdup; and second, that holdout should be of more concern than holdup to courts and other decision makers.

---

146. *See supra* notes 3–5 and accompanying text (citing the Supreme Court's 2006 decision providing four factors which courts should consider before determining to grant injunctive relief to patent owners and providing sources which survey relevant case law).

147. *See, e.g.*, Delrahim, The "New Madison" Approach to Antitrust and Intellectual Property Law, *supra* note 9, at 5 (arguing that, as a premise of the "New Madison" approach, "standard setting organizations should not become vehicles for concerted actions by market participants to skew conditions for patented technologies' incorporation into a standard in favor of implementers because this can reduce incentives to innovate and encourage patent hold-out"); Richard A. Epstein et al., *The FTC, IP, and SSOs: Government Hold-Up Replacing Private Coordination*, 8 J. COMPETITION L. & ECON. 1, 26–27 (2012) (noting that the "combined effect" of the FTC's prescriptions for when parties infringe would lead to an "inevitable slippage in the damage system"); F. Scott Kieff & Anne Layne-Farrar, *Incentive Effects From Different Approaches to Holdup Mitigation Surrounding Patent Remedies and Standard-Setting Organizations*, 9 J. COMPETITION L. & ECON. 1091, 1113 (2013) ("[I]nterpreting a RAND commitment as preventing patent holders from ever seeking an injunction" would allow infringers "to rationally consider the benefits of simply avoiding any up-front offer to take a license on any terms . . . .").

Commentators who express the view that holdout is a form of holdup base their argument on the fact that, in the standard-setting world in particular, firms compete to have their technologies included in a standard, knowing full well that only one of them will emerge victorious.[148] In such a winner-take-all market, the firm that prevails must receive an adequate return on its *ex ante* investment[149] in order to maintain the incentive for itself and other firms to participate in future competitions. *Ex post*, however, implementers can resort to delay and other tactics to resist paying adequate compensation.[150] On this telling, it is the technology owners whose irreversible, *ex ante*, sunk-cost investments have made them vulnerable to *ex post* appropriation.[151]

---

148. *See, e.g.*, Putnam, *supra* note 74, at 977 ("[S]tandardization is a winner-take-all process, in which—unlike market competition—there can be only one winner, by definition."); Delrahim, The "New Madison" Approach to Antitrust and Intellectual Property Law, *supra* note 9, at 8 ("If an inventor participates in a standard setting process and wins support for including a patented technology in a standard, that decision does not magically transform a lawful patent right into an unlawful monopoly.").

149. Actually, Putnam argues that in a winner-take-all market, adequate compensation must account for the *ex ante* probability that the firm would not prevail, since otherwise the industry-wide incentive will be too low. Putnam, *supra* note 74, at 980–81. Putnam's insight is similar in some respects to a critique two of us have made to the Baumol & Swanson auction model, but the resolution of these issues goes beyond the scope of this Article.

150. *See* Putnam, *supra* note 74, at 971 ("Just as innovators can hold up implementers by demanding a price that is 'too high' *ex post*, implementers can hold up innovators by demanding a price that is 'too low' *ex post*.").

151. *Id*. Note that Putnam is not arguing that firms are entitled to compensation simply because they have engaged in *ex ante* R&D. That would be a bad argument, since all a patent ever conveys is an opportunity to recoup, but if the market decides the invention has little or no commercial value the firm has no choice but to write off the R&D as a loss. Patents, in other words, reward success (building a better mousetrap, one that consumers want to buy), not effort (building a mousetrap that is no better than what is already on the market); and most analysts, including us, view this as a feature, not a bug. *See, e.g.*, Melamed & Shapiro, *supra* note 85, at 2118–20; Siebrasse & Cotter, *supra* note 82, at

1190–91. Nor do we understand him to be arguing that implementers are always wrong to resist patentees' demands. Empirical evidence indicates that over forty percent of patents litigated to judgment are invalidated in whole or in part, and that patentees overall win only about twenty-five percent of all cases litigated to judgment. *See, e.g.*, John R. Allison et al., *How Often Do Non-Practicing Entities Win Patent Suits?*, 32 BERKELEY TECH. L.J. 237, 269 (2017) (reporting, based on

At a high level, we acknowledge the parallel. Above, we defined holdup to be the ability of patent owners to extract royalties that exceed the *ex ante* baseline, because of sunk costs and opportunity costs incurred by the implementer.[152] The converse argument is that holdout allows patent users to pay royalties that are less than the *ex ante* baseline, because of sunk costs incurred by the patentee. At a more specific level, however, the analogy breaks down. In particular, it is often argued that denying injunctive relief is necessary to prevent holdup; the converse argument is that granting injunctive relief is necessary to prevent holdout.[153]

Here we disagree: that the problems of holdup and holdout are roughly analogous does not mean that the solutions must be analogous. Unlike the patent owner, who (if armed with an injunction) can force the implementer to switch or abandon the market altogether if the implementer is unwilling to agree to the owner's terms, implementers cannot simply refuse to deal with patent owners once they have begun using technologies covered by valid patents; at the very least, they would be liable for damages incurred through the date of judgment.[154] Moreover, if they *have*

---

an analysis of all patent infringement actions filed in federal district court in 2008–09, that "[t]he overall definitive patentee win rate in our data set for cases that went to a final judgment is 25.8%. This is consistent with prior work showing that patentees win approximately a quarter of their cases"); John R. Allison et al., *Our Divided Patent System*, 82 U. CHI. L. REV. 1073, 1099–1100, 1124–26 (2015) (reporting, based on the same data set, an overall invalidation rate of 42.6% for all patents litigated to judgment); Mark A. Lemley & Timothy Simcoe, *How Essential Are Standard-Essential Patents?*, 104 CORNELL L. REV. 607, 608 (2019) (finding that SEPs litigated to judgment are more likely to be valid, but (surprisingly) less likely to be infringed, than other patents). Invalidation and win rates in other countries are roughly comparable. For discussion, see THOMAS F. COTTER, COMPARATIVE PATENT REMEDIES: A LEGAL AND ECONOMIC ANALYSIS (2013).

152. *See supra* Part II.C.

153. *See supra* notes 3–5 and accompanying text.

154. The statement in the text may be a bit of an exaggeration. One could imagine a situation in which a powerful implementer simply *does* refuse to pay a patent owner whose lack of funds renders its ability to sustain an infringement suit illusory. Of course, if the patentee cannot afford to litigate at all, it does not matter whether the remedy is an injunction or not; but perhaps in some case a threat of injunctive relief would reduce the chance that the implementer would refuse to deal, because if the patentee did go to court, the implementer's exposure would be greater. We concede that under this set of circumstances, it may be

engaged in bad-faith behavior, for example, by dragging out negotiations for no reason other than to wear down the patent owner or gain some other advantage from delay, they make themselves vulnerable to a variety of additional sanctions. As Jorge Contreras notes, in such instances holdout "is simply willful patent infringement,"[155] which in the U.S. can result in the imposition of enhanced (up to treble) damages[156] and attorneys' fees.[157] (To be fair, though, such consequences are not guaranteed.)

Another possibility, in countries other than the United States, is that a court could award the infringer's profits, which typically will exceed a reasonable or FRAND royalty.[158] There also appears to be a general consensus, both in the U.S. and abroad, that even if a risk of patent holdup *generally* weighs against the imposition of injunctive relief, an injunction—and whatever leverage follows from its entry—may be appropriate when the implementer refuses to bargain in good faith.[159] Finally, if multiple implementers form

---

appropriate to consider the implementer as engaging in a form of opportunistic holdup by refusing to deal.

155.   Contreras, *supra* note 139, at 895.

156.   *See* Halo Elecs., Inc. v. Pulse Elecs., Inc., 136 S. Ct. 1923, 1930 (2016) ("This Court accordingly described § 284—consistent with the history of enhanced damages under the Patent Act—as providing that 'punitive or "increased" damages' could be recovered 'in a case of willful or bad-faith infringement.'") (citing Aro Mfg. Co. v. Convertible Top Replacement Co., 377 U.S. 476, 508 (1964)).

157.   *See* Octane Fitness, LLC v. Icon Health & Fitness, Inc., 572 U.S. 545, 553–54 (2014) (noting that district courts have discretion to award attorney's fees for "exceptional" cases in patent litigation).

158.   *See* Contreras et al., *supra* note 6, at 167–68, 191–92, 197–200 (providing examples of foreign jurisdictions awarding damages in excess of a FRAND royalty).

159.   *See, e.g.*, Case C-170/13, Huawei Techs. Co. v. ZTE Corp., 2015 E.C.J., ¶¶ 65, 71, https://perma.cc/5FYZ-DBA6 (last visited Nov. 3, 2019) (ruling that an SEP holder may only be found in breach of competition rules when it seeks an injunction against a potential licensee in certain circumstances); JAPAN PATENT OFFICE, GUIDE TO LICENSING NEGOTIATIONS INVOLVING STANDARD ESSENTIAL PATENTS 21 (2018), https://perma.cc/Q6YB-5SFU (PDF) ("Most courts have imposed limitations on . . . injunction[s] against implementers who have responded in good faith, and have determined that it would be appropriate for a rights holder . . . to seek an injunction when implementers have responded in bad faith during the negotiation process."); U.S. DEP'T OF JUSTICE & U.S. PATENT & TRADEMARK OFFICE, POLICY STATEMENT ON REMEDIES FOR STANDARDS-ESSENTIAL PATENTS SUBJECT TO VOLUNTARY F/RAND COMMITMENTS (2013),

an agreement to limit the royalties they will offer to SEP holders (which may be tantamount to a "buyer cartel"), then antitrust provides a potential means of attacking the holdout problem.[160]

That said, it remains possible that delay could result in harms that would not be fully remedied by such measures, if for example court-awarded damages are systematically under compensatory[161] or if legal doctrine declines to view certain harms as non-compensable.[162] This gets to the second point, that regardless of whether it makes sense to think of holdout as a form of holdup, it is a type of behavior that threatens the long-term health of the patent system and that policymakers therefore should not ignore. To that extent, we agree. Moreover, it is certainly possible that, as an empirical matter, the harms resulting from holdout outweigh the harms resulting from holdup, either in a particular case or industry, or in general.[163] In the absence of empirical evidence,

https://perma.cc/DF9S-ABW3 (describing appropriate determination of remedies in relation to patents bound by FRAND commitments); Motorola Mobility L.L.C., F.T.C., No. C-4410 (2013); Bosch GmbH, F.T.C., No. C-44377 (2012); Yabing Cui, *Across the Fault Lines: Chinese Judicial Approaches to Injunctions and SEP's*, CHINA IPR (June 5, 2018), https://perma.cc/V5LF-8YQP (last visited Nov. 3, 2019) (discussing the approach recently articulated in Chinese legal sources) (on file with the Washington and Lee Law Review).

160. For a discussion on this topic, see Peter C. Carstensen, *Buyer Cartels Versus Buying Groups: Legal Distinctions, Competitive Realities, and Antitrust Policy*, 1 WM. & MARY BUS. L. REV. 1 (2010).

161. As they are sometimes asserted to be, see, Einer Elhauge, *Do Patent Holdup and Royalty Stacking Lead to Systematically Excessive Royalties?*, 4 J. COMP. L. & ECON. 535, 557 (2008), though we're not convinced this is so, see Cotter, *supra* note 11, at 175, n.52.

162. In the U.S., for example, courts routinely award prejudgment interest and have discretion to compound it, but in many countries interest itself is discretionary or, if awarded, is not (or is rarely) compounded. For discussion, see Colleen V. Chien et al., *Enhanced Damages, Litigation Cost Recovery, and Interest, in* COMPLEX PRODUCTS, *supra* note 6, at 90, 111–14.

163. Arguably, the first-best response to holdout behavior enabled by deficiencies in the law of damages (e.g., refusing to award compound interest) would be to reform the law of damages. Until that happens, however, it probably is sensible for courts to take holdout risks into account when determining whether to grant injunctions. Moreover, if courts too readily deny injunctions, implementers may have an incentive to forgo voluntary bargaining even when such bargaining would be possible, the result being an increase in the error and adjudication costs that property rules avoid. Just because a risk of holdup may exist in a given case does not necessarily lead to the conclusion that courts should automatically refuse injunctions, to the detriment of all other considerations. *See generally* Richard A. Epstein & Kayvan B. Noroozi, *Why Incentives for Patent*

however, it would be foolish to ignore the risks of holdup just because the risks of holdout could be worse.

To repeat ourselves, if we are correct in concluding that the ability to engage in patent holdup depends on path dependence, that situations that are conducive to patent holdup are common, and that the three components of a holdup royalty can exist independently of one another, the risk of holdout provides no warrant for ignoring the risks of holdup, or vice versa. Put another way, even if one views holdout as a form of holdup, the question posed in any given case is whether one or the other or both are present; and if so, what remedy is appropriate to minimize the net harm. Injunctive relief may be an appropriate tool for fending off holdout in a given case, but it is not necessarily the best or only tool, depending on the circumstances; and in any event, courts should be aware of the risks resulting from both forms of behavior.

D. Other IP Doctrines That Mitigate Holdup Risks

Finally, although the point is not directly responsive to any of the above critiques, we would argue that several doctrines within IP law can be viewed as tools for reducing the risk of holdup under certain discrete circumstances, a fact that few observers up until now have explicitly noted. Note that we are *not* claiming that the reduction of holdup risk was a primary or even conscious goal of the courts and legislatures that developed these doctrines. Rather, we claim that, just as it sometimes argued (and disputed) that the common law evolved in a manner that tended to promote the goal of economic efficiency,[164] several doctrines in IP law can be interpreted as reducing holdup risks, albeit sometimes only imperfectly or secondarily to their main purpose.

Most obvious, perhaps, are the various equitable defenses and doctrines that arise in IP (particularly patent) litigation. For example, the doctrine of equitable estoppel precludes the patent

---

Holdout Threaten to Dismantle FRAND, and Why It Matters, 32 BERKELEY TECH. L.J. 1381 (2017).

164. For a brief overview of the efficient common law hypothesis, see, for example, RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW § 8.1 (9th ed. 2014).

owner from asserting what otherwise would be a valid claim of patent infringement if:

(1) "[t]he patentee, through misleading conduct, leads the alleged infringer to reasonably infer that the patentee does not intend to enforce its patent against the alleged infringer," (2) "[t]he alleged infringer relies on that conduct," and (3) "[d]ue to its reliance, the alleged infringer will be materially prejudiced if the patentee is allowed to proceed with its claim."[165]

By validating the defendant's reasonable reliance on the owner's implicit representation that it would not enforce its patent, the doctrine prevents the owner from extracting a holdup royalty *ex post*; and courts have applied the doctrine to, among other matters, cases involving patent ambush, which as discussed above[166] occurs when a firm misleads other members of an SSO into incorporating the firm's patented technology into a technical standard.[167] Other equitable doctrines that frequently come up in patent litigation—including laches,[168] prosecution laches,[169] prosecution history estoppel,[170] and the disclosure dedication rule[171]—play a similar role in backing up the implementer's

---

165. A.C. Aukerman Co. v. R.L. Chaides Constr. Co., 960 F.2d 1020, 1028 (Fed. Cir. 1992) (en banc), *abrogated in part on other grounds by* SCA Hygiene Prods. Aktiebolag v. First Quality Baby Prods., L.L.C., 137 S. Ct. 954 (2017).

166. *See supra* note 123–125 and accompanying text.

167. *See* Qualcomm, Inc. v. Broadcom Corp., 548 F.3d 1004, 1021–24 n.8 (Fed. Cir. 2008) (noting that equitable estoppel may generally be an appropriate legal framework for analysis of breaches of disclosure duties in the SSO context).

168. *See* Petrella v. Metro-Goldwyn-Mayer, Inc., 572 U.S. 663, 667 (2014) (holding that laches—"unreasonable, prejudicial delay in commencing suit"—do not bar a copyright owner's claim for damages brought within the statute of limitations, but may preclude equitable relief); *SCA Hygiene Prods. Aktiebolag*, 137 S. Ct. at 959 (extending *Petrella* to patent law).

169. *See* Cancer Research Tech. Ltd. v. Barr Labs., Inc., 625 F.3d 724, 728–32 (Fed Cir. 2010) (holding that a patent may be rendered unenforceable if the patent owner engaged in unreasonable, unexplained, prejudicial delay in prosecution).

170. *See* Festo Corp. v. Shoketsu Kinzoku Kogyo Kabushiki Co., 535 U.S. 722, 727 (2002):

> [W]hen the patentee responds to the rejection by narrowing his claims, this prosecution history estops him from later arguing that the subject matter covered by the original, broader claim was nothing more than an equivalent. Competitors may rely on the estoppel to ensure that their own devices will not be found to infringe by equivalence.

171. *See* PSC Computer Prods., Inc. v. Foxconn Int'l, Inc., 355 F.3d 1353, 1360 (Fed. Cir. 2004) ("[I]f one of ordinary skill in the art can understand the unclaimed

reasonable belief that it is safe to proceed with its product launch, and thus in forestalling holdup. Further, patent and copyright law offer protections to good-faith assignees or licensees of IP rights who lack notice of an earlier assignee's or licensee's rights;[172] and U.S. bankruptcy law shields the licensee of an IP right from having its license rejected by the trustee in bankruptcy, for the benefit of creditors.[173]

Second, in a recent paper titled *Trolls and Orphans*, Tun-Jen Chiang notes that owners of copyrighted "orphan works" can engage in a form of holdup similar to what PAEs are able to achieve when implementers are locked in to an *ex ante* technological choice.[174] Orphan works are works that were created or published recently enough that they are, or at least may be, still subject to

---

disclosed teaching upon reading the written description, the alternative matter disclosed has been dedicated to the public.").

172.   *See* 35 U.S.C. § 261 (2018):

> An interest that constitutes an assignment, grant or conveyance shall be void as against any subsequent purchaser or mortgagee for a valuable consideration, without notice, unless it is recorded in the Patent and Trademark Office within three months from its date or prior to the date of such subsequent purchase or mortgage.

See also 17 U.S.C. §§ 205(d), 205(e) (2018) (copyrights).

173.   *See* 11 U.S.C. § 365(n) (2018) ("If the trustee rejects an executory contract under which the debtor is a licensor of a right to intellectual property, the licensee under such contract may elect . . . to retain its rights . . . under such contract . . . ."); Mission Prod. Holdings, Inc. v. Tempnology, LLC, 139 S. Ct. 1652, 1657 (2019) (holding that a debtor-licensor's rejection of a trademark licensing contract in bankruptcy does not deprive the licensee of its right to use the mark); Jaffé v. Samsung Elecs. Co., 737 F.3d 14 (4th Cir. 2013) (holding that the application of 11 U.S.C. § 365(n) was necessary to ensure the protection of licensees under the foreign debtor's United States patents); Anthony Duggan & Norman Siebrasse, *The Protection of Intellectual Property Licenses in Insolvency: Lessons from the* Nortel *Case*, 4 PENN. ST. J.L. & INT'L AFF. 489 (2015) (comparing and contrasting protections granted to IP licensees in the U.S. and Canada using the bankruptcy-focused strategy developed in *Nortel*). These latter rules favoring good-faith licensees potentially could provide an analogy for efforts to ensure that FRAND commitments "run with the patent," that is, are enforceable against a subsequent assignee of a FRAND-committed patent. For discussion, see, for example, Jay P. Kesan & Carol M. Hayes, *FRAND's Forever: Standards, Patent Transfers, and Licensing Commitments*, 89 IND. L.J. 231 (2014).

174.   *See* Chiang, *supra* note 21, at 694 ("The orphan works problem occurs when potential users of a work fear the later emergence of an undiscovered copyright holder and therefore refrain from using the work.").

copyright protection.[175] Because the copyright term today is so long (consisting of the life of the author plus seventy years in many countries, including the United States),[176] it often can be difficult to determine whether a work is still subject to copyright at all, and if so who the owner is.[177]

As a result, contemporary writers, filmmakers, and other creative artists who want to reproduce or adapt works that were created several decades ago sometimes face a dilemma in deciding whether to devote resources to tracking down the work's copyright status and current owners, risk liability, or abandon the proposed use.[178] If the owner were to enjoin the use once it was underway, it would be able to extract a holdup royalty.[179] Among the possible solutions to this dilemma are for courts to award reasonable royalties only, in the event of litigation.[180] Alternatively, some commentators argue that courts should excuse the use altogether as a fair use, if the defendant has made reasonable (albeit unsuccessful) *ex ante* efforts to find the owner,[181] given that one of the standard rationales of fair use is to prevent transaction costs or other bargaining obstacles from derailing what would otherwise be productive uses of copyrighted works.[182]

---

175.   *Id.* at 697.

176.   *See* 17 U.S.C. § 302 (2018). For other countries, see Wikipedia's table of copyright durations by country, which to our knowledge is both comprehensive and accurate. *List of Countries' Copyright Lengths*, WIKIPEDIA, https://perma.cc/5UBZ-CL6B (last visited Nov. 3, 2019) (on file with the Washington and Lee Law Review).

177.   *See, e.g.*, REGISTER OF COPYRIGHTS, REPORT ON ORPHAN WORKS (2006), https://perma.cc/A99S-KD5X (PDF) (articulating possible solutions to the orphan works problem).

178.   *Id.* at 36.

179.   *Id.* at 37.

180.   *Id.*

181.   *See* Jennifer M. Urban, *How Fair Use Can Help Solve the Orphan Works Problem*, 27 BERKELEY TECH. L.J. 1379, 1397 (2012) (arguing that because the true owners of "orphan" works are unlikely to publicly release them, allowing their use through the fair use doctrine would "advance knowledge without undermining markets for the works"); *see also* Chiang, *supra* note 21, at 696 n.15 (proposing other possible solutions); David R. Hansen et al., *Solving the Orphan Works Problem for the United States*, 37 COLUM. J.L. & ARTS 1 (2013).

182.   *See, e.g.*, Wendy J. Gordon, *Fair Use as Market Failure: A Structural and Economic Analysis of the Betamax Case and Its Predecessors*, 82 COLUM. L. REV. 1600, 1614–15 (1982) (stating that "[a]n economic justification for depriving a copyright owner of his market entitlement only exists when the possibility of

A third example of how IP law engages with opportunism is the exhaustion doctrine, some version of which to applies to patents, copyrights, and trademarks.[183] In the patent law context, for example, although the owner of a valid patent generally has the right to exclude others from making, using, or selling a product incorporating the patented technology, once the owner authorizes the sale or other transfer of ownership of the product to another party, that party is free to use or dispose of that article as she sees fit; put another way, once the patent owner authorizes the sale of the patented article her right to control the further use or sale of that specific article is exhausted.[184] As Wenton Zheng observers, the exhaustion doctrine prevents patent owners from engaging in what we would describe as a form of classic holdup:

> . . . [I]magine a world in which there is not a rule of patent exhaustion, namely, a world in which a patentee is allowed to impose restrictions on the use or resale of a patented article without being required to inform the purchaser of the restrictions at the time of the sale of the patented article. In such a world, the purchaser faces an enormous amount of uncertainty as to what restrictions might be imposed by the patentee in the future[:] it could be a minor restriction that will not substantially hinder the purchaser's use of the patented article, or it could be a major restriction rendering the patented article entirely useless. This uncertainty makes it very difficult, if not impossible, for the purchaser to assess the value of the patented article. It is conceivable that if the uncertainty surrounding the rights that come with a purchase proves to be too much to bear, some risk-averse purchasers may forego the purchase altogether. Seen in this light, the doctrine of patent exhaustion could be understood as a way of preventing the seller from imposing after-the-fact restrictions so as to force the purchaser and the seller to bargain over the value of the patented article based on specific patent restrictions known to the purchaser at the time of the purchase.[185]

---

consensual bargain has broken down in some way.").

183. *See, e.g.*, Impressions Prods., Inc. v. Lexmark Int'l, Inc., 137 S. Ct. 1523, 1526 (2017) (holding that "[a] patentee's decision to sell a product exhausts all of its patent rights in that item, regardless of any restrictions the patentee purports to impose.").

184. *Id.*

185. Wentong Zheng, *Exhausting Patents*, 63 UCLA L. REV. 122, 154 (2016)

Framing the question in terms of opportunism and the parallel with holdup is helpful in assessing the proper scope of exhaustion. Because U.S. courts have interpreted the exhaustion doctrine to apply even when the seller of the product embodying the IP right puts the purchaser on notice of a post-sale restraint on subsequent use or sale,[186] the doctrine goes further than necessary to eliminate holdup risks, which would not exist if the purchaser was on notice *ex ante*.[187]

Finally, even patent law's novelty requirement,[188] which prevents anyone from claiming exclusive rights in subject matter that already is in the public domain, can be viewed as (in a sense) reducing the risk of some forms of opportunistic behavior (even though, as we again hasten to add, that is unlikely its principal or intended purpose). One could imagine a different system, under which (for example) a person who invests in commercializing an underutilized technology[189] or in proving its safety and effectiveness,[190] could acquire exclusive rights over that technology

---

(citations omitted).

186. *See, e.g.*, *Impression Prods.*, 137 S. Ct. at 1531, 1535 ("Once a patentee decides to sell—whether on its own or through a licensee—that sale exhausts its patent rights, regardless of any post-sale restrictions the patentee purports to impose, either directly or through a license.").

187. For an analysis of the aggregate welfare effects of the exhaustion doctrine, see Daniel J. Hemel & Lisa Larrimore Ouellette, *Trade and Tradeoffs: The Case of International Patent Exhaustion*, 116 COLUM. L. REV. ONLINE 17 (2016).

188. *See, e.g.*, 35 U.S.C. § 102(a)(1) (2018) (stating that "[a] person shall be entitled to a patent unless . . . the claimed invention was patented, described in a printed publication, or in public use, on sale, or otherwise available to the public before the effective filing date of the claimed invention"); In re Cruciferous Sprout Litig., 301 F.3d 1343, 1349 (Fed. Cir. 2002) (holding that a claimed invention lacks novelty if it is necessarily present—"inherent"—in the prior art). Independent invention nevertheless is not a defense against a claim of patent infringement, one possible reason being that (in theory) the defendant could have discovered the patent prior to launching its own product. A great deal of patent infringement probably is the result of independent invention, however, see generally Christopher A. Cotropia & Mark A. Lemley, *Copying in Patent Law*, 87 N.C. L. REV. 1421, 1451–57 (2009). Courts therefore still need to be careful in deciding upon the correct remedy for such infringement in order to avoid holdup risks.

189. *See, e.g.*, Ted Sichelman, *Commercializing Patents*, 62 STAN. L. REV. 341 (2010) (discussing how "[d]ecoupling the invention and commercialization functions of patent law" may substantially increase the commercialization of new technology).

190. *See* Benjamin N. Roin, *Unpatentable Drugs and the Standards of*

for a period of time. The patent system nevertheless prevents this from happening, and in doing so it affords others the freedom to access such subject matter without concern over being held up *ex post*.[191] On the other hand, if reducing holdup risks were the

---

Patentability, 87 TEX. L. REV. 503, 509–10 (2009). Note that someone who develops a new use of an existing product can obtain a method patent on that specific use, but not on the manufacture or other use of the product itself. *See* Arti K. Rai & Grant Rice, *Use Patents Can Be Useful: The Case of Rescued Drugs*, 6 SCI. TRANSLATIONAL MED. 248 (2014). Food and drug law, by contrast to patent law, sometimes does confer exclusive marketing rights on firms that have undertaken clinical studies of nonpatented drugs, in effect taking those drugs out of the public domain for a period of time. As a result, the *ex post* price of such drugs can increase substantially. For discussion, see COTTER, *supra* note 1, at 169–70 & n.125.

191.   To be sure, one might argue in response that if a technology was not being exploited *ex ante*, by definition there is no need to worry that exclusive rights would result in anyone being held up *ex post*. Conferring a blanket immunity on all uses of public domain subject matter nevertheless eliminates the risk that others, who may have already begun working on new applications of old inventions or who are unaware that the invention is no longer in the public domain, will be subject to holdup. One can still argue whether the immunity is, on balance, a wise response to what could be only a small risk of holdup, but our point is simply that conventional novelty doctrine presents on example by which conventional patent law forestalls a risk of *ex post* holdup.

        Further to the point, a variety of other doctrines in copyright and trademark law similarly function to keep public domain subject matter in the public domain, and thus to reduce holdup risks. *See, e.g.*, Dastar Corp. v. Twentieth Century Fox Film Corp., 539 U.S. 23, 33–38 (2003) (stating, as a general principle animating case law on preemption, functionality, and other IP doctrines, that "[t]he right to copy, and to copy without attribution, once a copyright has expired, like 'the right to make [an article whose patent has expired]—including the right to make it in precisely the shape it carried when patented—passes to the public'") (quoting *Sears, Roebuck & Co. v. Stiffel Co.*, 376 U.S. 225, 230 (1964)); *see also* 17 U.S.C. § 104A(d)(3) (2018) (permitting a "reliance party" to continue exploiting a derivative work based on a restored work and created before the date of copyright restoration, if the reliance party pays reasonable compensation where a copyright lapsed as a result of the author's noncompliance with copyright formalities, and has been restored pursuant to the Copyright Restoration Act); *id.* § 203(b)(1) (permitting the continued utilization of a derivative work prepared under authority of a grant of a transfer or license of copyright, prior to the termination of that transfer under § 203(a), in accordance with the terms of the grant); *id.* §§ 304(c)(6)(A), (d) (similar to § 203(b)(1), with respect to transfers effected prior to January 1, 1978). On the other hand, secondary liability doctrines such as contributory infringement sometimes enable patent and copyright law to assert claims against persons who sell non-IP-protected products that can be used to infringe patented or copyrighted articles. The logic of such doctrines, however, is that without a claim for secondary liability it sometimes would be impractical to enforce one's rights in the protected

primary concern, one might expect patent law to recognize independent invention as a defense, something that (for a variety of arguably valid reasons) it does not.[192] The analogy only goes so far.

V. Conclusion

Consistent with much of the economic and legal literature that preceded it, the analysis presented in this Article demonstrates that a holdup royalty can consist of three distinct components. More specifically, when an implementer commits to a particular technological path without clearing the necessary rights in advance, a patent owner who can credibly threaten an injunction can extract an *ex post* royalty that exceeds the *ex ante* baseline in an amount up to the sum of the implementer's sunk costs plus the opportunity costs of not having chosen the next-best nonpatented alternative *ex ante*. This latter cost in turn consists of both the *higher ex post* costs of redesign—not the *ex post* costs of redesign *as such*—if any, and the forgone benefits of not having chosen the alternative initially. Moreover, any of these three components—sunk costs, higher *ex post* redesign costs, and forgone benefits—can exist independently of the others, though it may be that all of them will be present to some degree in many instances, and all three share a common origin in path dependence.

The aim of this paper has been to disentangle these sources of holdup so that litigants and courts can better identify when holdup may be a risk, and when it is not. This is a first step in answering the ultimate question, such as whether an injunction should be granted to a successful patentee, not a last step. Even when the

---

article itself. *See* Thomas F. Cotter, *Four Questionable Rationales for the Patent Misuse Doctrine*, 12 MINN. J. L. SCI. & TECH. 457, 480–85 (2012) (noting this point, and also discussing commentary suggesting that the doctrine of patent and copyright misuse might work to prevent incursions upon the public domain).

192.   *See, e.g.*, COTTER, *supra* note 1, at 265 n.28 (2018) ("There are a variety of arguably plausible reasons for not exempting inadvertent infringement, among them that doing so might discourage firms from reading patents or increase the costs of administering and adjudicating patent disputes."). *See also* Shyamkrishna Balganesh, *Copyright and Good Faith Purchasers*, 104 CAL. L. REV. 269, 272 (2016) (noting that copyright law does not immunize the good-faith purchaser of a copy incorporating a copyrighted work from liability for the subsequent distribution of that copyrighted work).

evidence indicates that a serious risk of holdup may be present, it does not necessarily follow that courts should always deny an injunction to the prevailing patent owner. To be sure, if and when patent holdup occurs, it threatens substantial harms to static and dynamic efficiency, for reasons discussed in Part II above. Given the amount of money at stake in SEP cases in particular, the social welfare loss in any given instance could amount to billions of dollars. The questions are difficult, and ultimately empirical, ones. We outline them here, not to take a position on them, but only to emphasize that our analysis of the sources of holdup is only one piece of the puzzle. Nonetheless, it is an important piece of the puzzle. For example, if the costs of switching to a new technology are very high, but the evidence shows that the costs of adopting it would have been just as high *ex ante*, then our analysis suggests (barring problems of sunk costs or differential profits), that the question of whether to grant an injunction is easier than it might seem.

Finally, each legal system must decide for itself how, if at all, courts should address holdup risks. Some may decide that it's best to make case-by-case determinations (perhaps along the lines of the examples in Part III.B), while others, conceivably, may conclude that bright-line rules (e.g., no injunctions in PAE or SEP cases absent proof of holdout) would conserve on adjudication costs at a tolerable risk of error.[193] Moreover, given the differences that exist among individual countries' legal systems and institutions, nations must decide for themselves whether holdup risks, if they are to be addressed at all, are best handled by the law of patent remedies or contract, as in the U.S., or by competition law, the civil law doctrine of abuse of right, or something else.[194]

---

193. In the U.S., the Supreme Court has warned against the use of bright lines, or even presumptions, in applying *eBay*. *See* eBay, Inc. v. MercExchange, L.L.C., 547 U.S. 388, 392–95 (2006) (emphasizing "traditional equitable principles" over "categorical rule[s]"). In other common law countries, by contrast, there is usually a presumption in favor of injunctive relief, which the infringer can try to rebut. *See* Siebrasse et al., *supra* note 7, at 154–55. On the pros and cons of rules versus standards generally as applied to patent damages issues, see Cotter, *supra* note 41.

194. *See generally* Thomas F. Cotter, *Comparative Law and Economics of Standard-Essential Patents and FRAND Royalties*, 22 Tex. Intell. Prop. L.J. 311 (2014) (discussing potential solutions to the patent holdup problem).

However courts and other decision makers throughout the world choose to address these matters, one thing is certain: before crafting a legal response to an alleged problem, it's good to know exactly what the problem is. We hope this paper will enable policymakers and their advisors to better understand what patent holdup is, and will spur them to develop better methods for detecting when a holdup threat is present—and when it's not.

VI. Technical Appendix

We presented the basics of the model in the text, but here are a couple of finer points.

Standards and Network Value

Many authors have highlighted the acute risk for holdup to occur in the standard-setting context.[195] The principal value of standard-setting is that it helps to ensure interoperability among a large number of complementary technologies. But, *ex post*, an implementer enjoys this value only so long as it adheres to the standard adopted *ex ante*. Our model makes it easy to see how this creates a pervasive holdup threat (although we do not claim this is the only thing that contributes to holdup in the standards context).[196]

Suppose there is a positive mass η of implementers, and let η_t denote the measure of those that commit to technology t at the ex ante stage, where $\eta_1 + \eta_2 = \eta$. We assume that the marginal profit an implementer gets from t is a strictly increasing function of η_t, which we denote $v_t(\eta_t)$. This reflects the network value of interoperability. To isolate focus to these network effects, assume for simplicity there are no fixed costs associated with either technology ($K_t = 0$, each t).

The implementers coordinate the choice of technologies through a standard-setting organization to ensure that, whatever their choice, they will enjoy the full network value. That is, they

---

195. *See, e.g.*, Farrell et al., *supra* note 21, at 603 (discussing holdup risks).

196. For example, it could be also be that sunk costs tend to be larger in this context. Such possibilities are similarly easy to express using our model.

ensure that $\eta_{t^*} = \eta$ for some selection t^*. As such, at the *ex ante* stage, the firms compare $\pi_1 = v_1(\eta)$ to $\pi_2 = v_2(\eta)$. Assume this is larger for technology 1. Then, if an implementer switches unilaterally to technology 2 *ex post*, it anticipates earning a profit of $\pi_{2|1} = v_2(0)$.

Hence, if the patentee asserts its rights against an individual implementer, the holdup rents are $\mathcal{H} = \mathcal{N}_2$, where $\mathcal{N}_2 \equiv v_2(\eta) - v_2(0)$ is the forgone network value of technology 2. This is the opportunity cost of adopting technology 2 *ex post* rather than *ex ante*. While it may be intuitive to think that holdup serves to extract the network value of the adopted standard—technology 1—that is not the case. It is not the forgone network value associated with technology 1 that constitutes the holdup rents, but rather the forgone network value associated with *not* having selected technology 2 *ex ante*.

Uncertainty

Now we extend the model to account for patent uncertainty, and for bargaining in the shadow of litigation. Let θ ($0 \le \theta \le 1$) denote the probability that the patent would be held valid and infringed by the implementer's use of technology 1. (For expositional simplicity, we will just refer to "validity" and not "validity and infringement.") We now let $\mathcal{F}_\theta^A$ denote the maximal fee negotiated ex ante, conditional on θ, and similarly for $\mathcal{F}_\theta^P$. Note that the term cF^A defined previously can be expressed as $\mathcal{F}_1^A$, since we assumed away patent uncertainty. Note further that the fee will be zero if the patent is known to be invalid with certainty ($\mathcal{F}_0^A = 0$). As Lemley and Shapiro note (in a similar framework),[197] when we allow for any probability θ, an intuitive benchmark for the maximal *ex ante* fee is

$$\mathcal{F}_\theta^* = \theta \mathcal{F}_1^A + (1 - \theta)\mathcal{F}_0^A$$

---

197. *See* Lemley & Shapiro, *supra* note 11, at 2001 (discussing benchmarks for reasonable royalties).

$$= \theta \mathcal{F}_1^A$$

This is the expected value of the maximal fee that would be charged if the patent's validity were somehow revealed *ex ante*. Now suppose that, at the *ex ante* stage, the parties bargain in the shadow of litigation. We assume litigation costs are zero for simplicity. If the patentee wins, the court will issue a remedy whose monetary value is R. This is either a damages award or the forgone profits if an injunction is issued and enforced (forcing the implementer to switch to the alternative). We interpret a value $R > \mathcal{F}_1^*$ as excessive, since the remedy is more costly than the largest amount the implementer would pay conditional on knowing the patent is valid and infringed. If firm 1 adopts technology 1 *ex ante* and the parties litigate, then the implementer's expected total profits are

$$\theta(\pi_1 - K_1 - R) + (1 - \theta)(\pi_1 - K_1) = \pi_1 - K_1 - \theta R$$

Then, if *ex ante* negotiations break down, the implementer will either adopt the alternative, or else risk litigation. In particular, it will adopt the alternative if and only if $\pi_2 - k_2 > \pi_1 - k_1 - \theta R$. Note that if the implementer prefers risking litigation to adopting the alternative, then the maximal *ex ante* fee is θR. Otherwise the fee is $\pi_1 - k_1 - (\pi_2 - k_2) = \mathcal{F}_1^A$. This means that the maximal *ex ante* fee becomes

$$\mathcal{F}_\theta^A = \min\{\mathcal{F}_1^A, \theta R\}$$

The top and bottom cases refer to adoption of the alternative and deliberate infringement, respectively. Note that this implies that even *ex ante* fees are excessive in the sense that $\mathcal{F}_\theta^A > \mathcal{F}_\theta^*$; this occurs when remedies are excessive ($R > F_1^*$). In this case, there is some amount of systematic overcharging that is unrelated to holdup. This was one of the important results in Lemley and Shapiro.[198]

---

198.   *Id.* at 2003–04.

Now consider the *ex post* stage, after the implementer has committed to the patented technology. If the implementer still decides to infringe and risk losing in court, its expected profits at the margin will be $\pi_1 - \theta R$. But if it switches to the alternative (thereby avoiding liability), its profits will be $\pi_{2|1} - k_{2|1}$. The latter is the more profitable option if and only if $\theta R > \mathcal{F}_1^P$. Thus, given any probability θ, the maximal *ex post* fee will be

$$\mathcal{F}_\theta^P = \begin{cases} \mathcal{F}_1^P, & \text{if } \theta R > \mathcal{F}_1^P \\ \theta R, & \text{otherwise.} \end{cases}$$

Similar to the *ex ante* case, the fee is always excessive ($\mathcal{F}_\theta^P > \mathcal{F}_\theta^*$) when remedies are excessive. When $\theta R > \mathcal{F}_1^A$, so that the implementer prefers the alternative to deliberate infringement, things work out just like in our simpler model. In this case, it is easy to see that the holdup rents are $\mathcal{H} = \mathcal{F}_1^P - \mathcal{F}_1^A = k_1 + \Delta k_2 + \Delta \pi_2$, which is exactly the same as before (just with added subscripts). In this case, it is still true that holdup occurs if and only if there is path dependence. And note that, if remedies are excessive, then there is systemic overcharge *in addition to* holdup. No holdup occurs if the implementer prefers litigation to the alternative, for in both *ex ante* and *ex post* bargaining, the fee is θ.