



Content downloaded/printed from

[HeinOnline](#)

Mon Aug 19 19:07:50 2019

Citations:

Bluebook 20th ed.

Erik Hovenkamp, Antitrust Law and Patent Settlement Design, 32 HARV. J. L. & TECH. 417, 478 (2019).

APA 6th ed.

Hovenkamp, E. (2019). Antitrust law and patent settlement design. Harvard Journal of Law Technology (Harvard JOLT), 32(2), 417-478.

Chicago 7th ed.

Erik Hovenkamp, "Antitrust Law and Patent Settlement Design," Harvard Journal of Law & Technology (Harvard JOLT) 32, no. 2 (Spring 2019): 417-478

McGill Guide 9th ed.

Erik Hovenkamp, "Antitrust Law and Patent Settlement Design" (2019) 32:2 Harv JL & Tech 417.

MLA 8th ed.

Hovenkamp, Erik. "Antitrust Law and Patent Settlement Design." Harvard Journal of Law & Technology (Harvard JOLT), vol. 32, no. 2, Spring 2019, pp. 417-478. HeinOnline.

OSCOLA 4th ed.

Erik Hovenkamp, 'Antitrust Law and Patent Settlement Design' (2019) 32 HARV J L & TECH 417

Provided by:

USC Law Library

-- Your use of this HeinOnline PDF indicates your acceptance of HeinOnline's Terms and Conditions of the license agreement available at <https://heinonline.org/HOL/License>

-- The search text of this PDF is generated from uncorrected OCR text.

-- To obtain permission to use this article beyond the scope of your license, please use:

[Copyright Information](#)

Use QR Code reader to send PDF to your smartphone or tablet device



ANTITRUST LAW AND PATENT SETTLEMENT DESIGN

Erik Hovenkamp*

TABLE OF CONTENTS

I. INTRODUCTION.....	418
II. PATENT SETTLEMENTS BETWEEN RIVALS.....	424
<i>A. Horizontal Restraints: Examples</i>	424
<i>B. What's the Problem?</i>	426
1. Contrast with Settlements in Other Contexts	427
<i>C. The Scope of the Patent Doctrine</i>	429
<i>D. Actavis and Proportional Effects</i>	431
1. Pay-for-Delay Settlements	433
2. The Actavis Decision.....	435
<i>E. The Administrability Problem</i>	436
III. THE ECONOMICS OF SETTLEMENT DESIGN: AN OVERVIEW	438
<i>A. Relevant Aspects of Settlement Design</i>	439
<i>B. Licensing Restraints</i>	440
1. Applications	445
<i>C. Reverse Payments and Counter-Restraints</i>	450
1. Reverse Payments	451
2. Counter-Restraints.....	454
3. The Pareto Problem	456
IV. MULTIPLE ENTRY, HATCH-WAXMAN, AND OTHER	
CONSIDERATIONS.....	457
<i>A. Clarifying the Hatch-Waxman Problem</i>	460
<i>B. Litigation Costs</i>	462
<i>C. Cross-Licensing Settlements</i>	463
V. REFRAMING THE ANTITRUST QUESTION.....	465
<i>A. Administration</i>	465
<i>B. Demystifying Causation</i>	471
<i>C. Evaluating Potential Objections</i>	473
1. What if the Patentee is Risk-Averse?	473
2. Appeals to Non-Pareto Hypotheticals	475
VI. CONCLUSION	477

I. INTRODUCTION

Antitrust usually prohibits rival firms from striking agreements that forestall competition. Patent settlements provide an exception, however, because a patent on a significant technology may provide a lawful basis for excluding competitors from the marketplace. Problematically, firms always prefer to restrain competition into monopoly (and share in the proceeds), even if they privately believe that the patent is very likely invalid or noninfringed.¹ This might not be such a challenging problem if it were easy to discern whether a given patent is valid and infringed. But in practice this is almost never the case.²

The settlement problem is emblematic of the analytical difficulties that emerge at the intersection of antitrust and patent law, which has long been a source of widespread confusion and debate.³ In some cases, the settlement restrains inter-party competition in a way that is not even facially authorized by patent law, even if the relevant patent is definitively valid and infringed.⁴ But in the more challenging cases — which are the focus of this paper — the firms' settlement restrains competition in a manner that potentially falls within the patentee's exclusionary entitlement. Specifically, the settlement forecloses competition by no more than a permanent injunction would, which is the most restrictive remedy available for patent infringement.⁵ For example, if a settlement limits the monthly sales of a patented product by a rival licensee, it

* Postdoctoral Fellow, Harvard Law School, Project on the Foundations of Private Law; Visiting Fellow, Yale Law School, Information Society Project. I am grateful to David Abrams, Janet Freilich, Kyle Rozema, and Christopher Yoo for providing helpful comments.

1. See, e.g., Daniel A. Crane, *Ease over Accuracy in Assessing Patent Settlements*, 88 MINN. L. REV. 698 (2004); Michael J. Meurer, *The Settlement of Patent Litigation*, 20 RAND J. ECON. 77 (1989); George L. Priest, *Cartels and Patent License Arrangements*, 20 J.L. & ECON. 309 (1977); Carl Shapiro, *Antitrust Limits to Patent Settlements*, 34 RAND J. ECON. 391 (2003).

2. See, e.g., *In re Cipro Cases I & II*, 348 P.3d 845, 859 (Cal. 2015) (“[P]atents are in a sense probabilistic . . . they grant their holders a potential but not a certain right to exclude.”); Ian Ayres & Paul Klemperer, *Limiting Patentees' Market Power Without Reducing Innovation Incentives: The Perverse Benefits of Uncertainty and Non-Injunctive Remedies*, 97 MICH. L. REV. 985 (1999); Mark A. Lemley & Carl Shapiro, *Probabilistic Patents*, 19 J. ECON. PERSPECTIVES 75 (2005).

3. See Louis Kaplow, *The Patent-Antitrust Intersection: A Reappraisal*, 97 HARV. L. REV. 1815, 1815 (1984) (“The intersection of antitrust law and patent policy has proved to be a source of perpetual confusion and controversy since the passage of the Sherman Act nearly a century ago.”); see also, e.g., WARD S. BOWMAN, JR., *PATENT AND ANTITRUST LAW: A LEGAL AND ECONOMIC APPRAISAL*, at xi–xii (1973); William F. Baxter, *Legal Restrictions on Exploitation of the Patent Monopoly: An Economic Analysis*, 76 YALE L.J. 267, 277 (1966).

4. For example, when rival patent holders fix prices of patented products, this is plainly anticompetitive; the Patent Act does not authorize restraints on competition between separately patented technologies. See, e.g., *United States v. Line Material Co.*, 333 U.S. 287, 307–08 (1948).

5. 35 U.S.C. § 283 (2012) (authorizing injunctive relief “in accordance with the principles of equity, or on such terms as the court deems reasonable”).

clearly diminishes competition. Indeed, if not for the patent, it would be illegal *per se*. But it is less restrictive than an injunction, which would not permit the rival to make any sales at all.

As outlined below, antitrust analysis of these settlements has become increasingly concerned with how patent litigation would have played out but for the settlement. This is now widely-regarded as the appropriate benchmark for evaluating the settlement's competitive effects.⁶ To that end, scholars and courts have framed the antitrust inquiry to require evidence or "signals" bearing on the likelihood that the relevant patent is invalid or noninfringed. This gives the antitrust claim a "case-within-a-case" structure, as liability then hinges on speculation about the expected outcome of counterfactual patent litigation. Notwithstanding the challenges this presents in practice, such speculation is thought to be logically necessary to answer the antitrust question.

But this is not so. In fact, it is the *design* of the settlement agreement — the way it restrains competition or otherwise influences the distribution of profits — that determines how the parties' ultimate agreement will compare with their expectations about litigation.⁷ As I argue below, this supports a wholesale shift in how the antitrust inquiry is conceptualized and administered. In short: evaluate the agreement, not the patent.

Consider an example illustrating the underlying problem. A patent holder and its only rival are very confident that the rival's product does not infringe the operative patent — so confident that the rival will not agree to pay even a moderate royalty, as it gives little weight to the litigation threat. But the patentee proposes an alternative settlement, which restrains the rival in a different way. It offers to license the rival to be the exclusive seller of the patented product in one half of the country, so long as it does not make sales elsewhere;⁸ the patentee will then be the exclusive seller in the other half of the country.⁹ This generates a monopoly, albeit a divided one, since there is no competition in any given territory. It is thus easy to see why the rival accepted this proposal, as the settlement essentially creates a cartel.¹⁰ Still, the arrangement is no more restrictive than an injunction would have been.

6. See *infra* Section II.D.

7. The economic theory underpinning this general point was developed in a recent economics article by Jorge Lemus and myself. Erik Hovenkamp & Jorge Lemus, Proportional Restraints in Horizontal Patent Settlements (Jan. 17, 2019) (unpublished manuscript), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3259249 [<https://perma.cc/4Y2H-S2GH>].

8. The Patent Act authorizes a patentee to grant a territorially limited license, which may be an exclusive license. 35 U.S.C. § 261.

9. Absent the patent, this would be naked market division, which is illegal *per se*. See, e.g., *Palmer v. BRG of Georgia, Inc.*, 498 U.S. 46, 49 (1990).

10. Total profits decline as a market grows more competitive. Hence, under an evenly divided monopoly (a kind of cartel), each firm earns more than it would under open competition.

The problem is that litigation would have produced a much more competitive result — not with certainty, but in expectation.¹¹ The patentee would likely have lost, in which case the rival would have an unambiguous right to compete freely. Framing the antitrust inquiry to focus on the expected result of litigation is relatively new. Considerations of counterfactual litigation were rarely raised, let alone dispositive, in cases prior to the 21st century.¹² In a 2003 article, Carl Shapiro proposed that all horizontal settlements should be evaluated based on how their competitive effects compare to the expected result of litigation.¹³ This approach quickly came to dominate the antitrust literature¹⁴ and was recently embraced by the Supreme Court in *FTC v. Actavis*.¹⁵ Under this standard, which I call “the proportional-effects rule,” antitrust’s goal is for settlements to be *proportional*: to restrain competition to an extent commensurate with the expected result of patent litigation.¹⁶

At the outset, this standard would appear to suffer from a major administrability problem.¹⁷ How do you enforce a standard that hinges on the likelihood that a patent would have been held valid and infringed in counterfactual litigation? This would ostensibly require a case-within-a-case analysis of the relevant patent. However, in one well-known settlement format — “pay-for-delay,” also known as a “reverse payment” settlement — many antitrust commentators advocate a workaround. In such settlements, the patentee makes a lump sum “reverse payment” to its prospective rival in exchange for the latter’s agreement to delay its entry until later in the patent term, often soon before expiration.¹⁸ A large reverse payment is thought to be a suitable proxy for

11. For example, suppose the patent is 10% likely to be valid and infringed, and that total profits under open competition are \$50M, whereas the monopoly profit is \$100M. Then, in expectation, litigation would result in a total profit of $90\% \times \$50M + 10\% \times \$100M = \$55M$. That the settlement preserves the full \$100M reflects that it is far less competitive than the expected result of litigation.

12. *Valley Drug* was one of the earliest such cases. *Valley Drug Co. v. Geneva Pharm., Inc.*, 344 F.3d 1294, 1312 (11th Cir. 2003) (“[The] agreement should be considered in light of the likelihood of [the patentee] . . . obtaining [an injunction].”).

13. Shapiro, *supra* note 1.

14. See *infra* Part III.

15. 570 U.S. 136, 157 (2013) (concluding that “the relevant anticompetitive harm” arises where a settlement acts to “prevent the risk of competition”) (emphasis added).

16. See *infra* Section II.D.

17. See *infra* Section II.E.

18. There is extensive literature on these settlements. See, e.g., Michael A. Carrier, *Unsettling Drug Patent Settlements: A Framework for Presumptive Illegality*, 108 MICH. L. REV. 37 (2009); Daniel A. Crane, *Exit Payments in Settlement of Patent Infringement Lawsuits*, 54 FLA. L. REV. 747 (2002); Aaron Edlin et al., *The Actavis Inference*, 67 RUTGERS UNIV. L. REV. 585 (2015); Einer Elhauge & Alex Krueger, *Solving the Patent Settlement Puzzle*, 91 TEX. L. REV. 283 (2012); Scott C. Hemphill, *Paying for Delay: Pharmaceutical Patent Settlements as a Regulatory Design Problem*, 81 N.Y.U. L. REV. 1553 (2006).

the patent's likely invalidity.¹⁹ The Supreme Court's *Actavis* decision endorsed this proxy-based approach,²⁰ which has since been dubbed the "Actavis Inference."²¹

However, there remain major obstacles to enforcement of the proportional-effects rule. First, the Actavis Inference is applicable only in settlements that involve reverse payments. Many settlement arrangements — including essentially all types that received antitrust scrutiny prior to the 21st century — do not contain such payments. Second, even in post-*Actavis* pay-for-delay cases, some courts have essentially reverted to the requirement that a private plaintiff show (directly, not by proxy) that the patent is invalid.²² In practice, this may be an insurmountable barrier to enforcement.²³

These problems could be avoided entirely if the patent issues were wholly removed from the antitrust analysis. In a recent economics paper, Jorge Lemus and I demonstrate that a settlement's proportionality (or lack thereof) is ultimately driven by its design — the particular way it restrains competition, along with any other provisions influencing the distribution of profits.²⁴ That paper provides a formal economic theory of the relationship between settlement design and proportionality. The theory relies on only the most general properties that all restraints exhibit and can be applied within any economic model of oligopoly competition.

The present article applies those economic insights to antitrust law and practice. Among other things, I emphasize that: (a) this approach can be administered much more practicably, accurately, and broadly than the prevailing patent-focused approach;²⁵ (b) it is consistent with the Supreme Court's *Actavis* decision, notwithstanding that it removes the patent issues from the antitrust claim;²⁶ (c) it simplifies the antitrust analysis by disentangling the relevant antitrust violation from the extent of the resulting harm;²⁷ and (d) it clarifies a number of critical errors in arguments advocating against antitrust intervention in patent settlements.²⁸

19. See, e.g., Shapiro, *supra* note 1, at 408 ("Presumably, the patent holder would not pay more than avoided litigation costs unless it believed that it was buying a later entry than it expects to face through the litigation alternative.").

20. *Actavis*, 570 U.S. at 157–58.

21. Edlin et al., *supra* note 18, at 585.

22. See *infra* Section V.B.

23. See, e.g., Kevin B. Soter, *Causation in Reverse Payment Antitrust Claims*, 70 STAN. L. REV. 1295, 1320 (2018) (discussing approach taken by some courts wherein "plaintiffs who cannot produce sufficient proof (distinct from the [Actavis Inference]) about patent validity are bound to fail as a matter of law.").

24. Hovenkamp & Lemus, *supra* note 7. For a detailed breakdown of the relevant aspects of settlement design, see *infra* notes 123–132 and accompanying text.

25. See *infra* Section V.A.

26. See *infra* Section II.D.

27. See *infra* Section V.B.

28. See *infra* Section V.C.

In broad outline, a settlement's design determines what the parties can agree on in relation to their litigation expectations. That is precisely the comparison that drives antitrust liability under the proportional-effects rule. This leads to a counterintuitive result: the "design-focused" approach makes it possible to administer the proportional-effects rule — a standard intimately concerned with probability — without having to assess the probability that any particular patent is valid and infringed. This brings the antitrust analysis back into familiar territory, for it can focus entirely on the nature of the agreement; no more case-within-a-case. And, significantly, this analysis can be applied to *all* kinds of settlements, including those involving multiple rivals or multiple patentees.²⁹

In fact, one example of this design-focused approach in action is already well-known. In a "pure delay" agreement, the patentee's rival agrees that its entry will be delayed for some negotiated period of time, but it does not receive any reverse payment. As numerous commentators have noted,³⁰ this settlement forces the firms to settle on terms that emulate their litigation expectations no matter what they happen to be. For instance, suppose there are ten years remaining in a patent term, and that the firms think the patent has a 30% chance of being held valid and infringed, in which case the patentee would obtain an injunction. Then the patentee's rival will agree to a delay period of up to three years, whereas the patentee will agree to a delay period of no less than three years. Thus, they can only mutually agree on a three-year delay period, which is proportional. If the litigation odds were instead 75%, then they could only agree on 7.5 years, which is again proportional, and so on for other possible odds.³¹

Note that in working through the pure delay example, we did not ask about the "true" likelihood of validity and infringement, and by implication we had no reason to consider validity proxies. Indeed, the analysis was not even framed in a way that contemplates a particular case or patent. Instead, we allowed for any odds. For each possible set of odds, we asked what the firms could agree on, which depends not just upon the odds, but also their pure delay settlement design, since that determines what profits they will end up with when they restrain competition to any particular extent. Then, in each case, we simply asked if the outcome was proportional. And, in this case, it always was. The key economic point, whose policy relevance is this paper's focus,

29. See *infra* Part IV.

30. See, e.g., Hemphill, *supra* note 18, at 1588–89.

31. In practice, most delay-entry settlements occur in the pharmaceutical industry, which is subject to regulations creating incentive problems that can skew bargaining outcomes away from the proportional result. But, as explained in Part IV, this is ultimately a problem with the regulations themselves, for exactly the same problem arises no matter what settlement format the firms use.

is that we can evaluate every type of settlement using this general approach. For example, as noted in Part III, if the firms had relied on a price or output restraint rather than delayed entry, we would find that they always end up agreeing on terms that restrain competition disproportionately — they simply cannot agree on the proportional outcome. By contrast, a simple per-unit royalty is generally proportional, just like pure delay.

Under this approach, the focus is on how a settlement's design affects private bargaining in the shadow of patent litigation. A problematic settlement design has the property that, no matter the odds of counterfactual patent litigation, private bargaining always leads to disproportionate anticompetitive effects. This is either because the firms cannot agree on anything else, or because they both stand to gain from excessively restraining competition. The result is that we need not assess the odds of validity or infringement to know that the settlement will restrain competition disproportionately. Instead, such odds may become relevant only at the remedy stage, namely in computing damages.³²

By contrast, under a socially desirable settlement format, the agreed-upon competitive effects are always commensurate with the firms' litigation expectations. This is not because the firms do not want larger profits — of course they do. Instead, this happens because the settlement design makes it impossible to effect a disproportionate result without the arrangement becoming unacceptable to the patentee's rival. When the firms settle in such a way, there is no need to further scrutinize the patent — by proxy or otherwise — to assess the settlement's antitrust compliance. In fact, one can interpret a proportional settlement format as simply putting the firms in the same position as ordinary settling litigants outside the antitrust-IP context. In other areas of private law, the parties can only mutually agree on terms that comport with their litigation expectations.³³

The balance of the paper is organized as follows: Part II clarifies the underlying antitrust problem and describes how courts and commenters have converged toward the proportional-effects rule over the last two decades. Part III provides a nontechnical overview of the economics of settlement design, with emphasis on simple examples and links to the case law. Part IV considers a number of important extensions, such as settlements with a multiplicity of rivals, and the impact of Hatch-Waxman regulations on settlements involving pharmaceutical patents, which is the context generally associated with pay-for-delay.

32. See *infra* notes 277–278 and accompanying text.

33. See *infra* Section II.B.

Part V addresses how the design-focused approach should be administered in practice. This part also considers and responds to potential objections, after which the paper concludes.

II. PATENT SETTLEMENTS BETWEEN RIVALS

The antitrust case law contains a vast and diverse set of restrictive settlements among patent holders and their product market competitors. These agreements restrain competition in all kinds of different ways and facilitate ways to ensure that each party secures an acceptable payout. The next part illustrates this with a number of examples from the case law. However, it is helpful to begin by breaking down the underlying antitrust problem. Throughout this paper, we focus on situations in which (1) the patent holder has market power; and (2) the patent, if upheld as valid and infringed, would act as a significant barrier to the rival's ability to compete.³⁴ For brevity, I will frequently refer to validity alone, rather than validity and infringement.

All kinds of restraints inhibit the rival's ability to compete with the patent holder, causing it to make fewer sales over the operative time period.³⁵ This leaves the market less competitive than if the patent were deemed invalid or noninfringed or if the rival had purchased a license with a lump-sum fee, which would not restrain competition.³⁶ The results are twofold. First, total profits rise due to the reduction in competition.³⁷ Second, the challenger is put at a competitive disadvantage, leaving it with a smaller share of total profits. This must always benefit the patent holder, as it receives a larger share of a larger pie. By contrast, the rival's profits will typically fall,³⁸ albeit by less than the patentee's rise. For example, the patent holder's profits might increase by \$2M, while the rival's fall by \$1M.

A. Horizontal Restraints: Examples

One reason for the longstanding antitrust confusion over patent settlements is that they are so varied, involving all kinds of different restraints on competition. Broadly, a restraint is an arrangement that

34. If these conditions are not satisfied, then restraining the rival's use of the patented invention will not have a significant impact on competition, in which case the restraint is unlikely to raise serious antitrust concerns.

35. A settlement's restrictions can persist no longer than the balance of the patent term. To go any farther would be transparently unlawful.

36. A lump sum license fee is a fixed payment by the rival to the patent-holder that, once executed, leaves the rival with an unrestricted right to use the invention. This does not suppress competition; it merely reallocates the total profits that accrue from open competition.

37. Competition erodes total profits by inducing lower prices than a monopolist (or a cartel) would set.

38. There are exceptions in which the restraint can actually benefit the rival, relative to the absence of a restraint. See *infra* notes 153–163 and accompanying text.

limits a firm's ability to engage in some commercial activity that is germane to the competitive process or that makes it costlier to engage in such activity. This paper's focus is on "horizontal restraints," or those arising from an agreement between competitors.³⁹

Consider some examples. One familiar restraint is a per-unit royalty imposed on the rival licensee's sales. From the rival's perspective, this is no different from an increase in marginal cost, which naturally induces the rival to set a higher price and make fewer sales.⁴⁰ Alternatively, a settlement may restrain the rival's entry in some way. For instance, in pay-for-delay settlements, discussed earlier, the relevant restraint is a delay on entry. But the pay-for-delay format also includes a reverse payment.⁴¹ An entry restraint could be territorial, limiting where the rival-licensee can operate.⁴² Yet another form of entry restriction is a field-of-use restraint, which limits what kinds of patented products the rival-licensee can sell.⁴³

The antitrust concerns may be somewhat more acute in cases where the restraint applies to the rival's output or price.⁴⁴ In some cases, the patentee may agree to a restraint on its own use of the commercial invention, in addition to the restraint on its rival.⁴⁵ A later section discusses such "counter-restraints."⁴⁶ Additional complications arise in cross-licensing settlements between competing patentees. In some such

39. By contrast, vertical restraints are those arising between firms in a buyer-seller relationship, such as a manufacturer and its distributor. Although there are important exceptions, most such restraints do not raise serious antitrust concerns. *See* *Leegin Creative Leather Prods., Inc. v. PSKS, Inc.*, 551 U.S. 877, 882 (2007) (overturning prior holding that vertical price restraints are illegal *per se*).

40. *See, e.g.*, *Applera Corp. v. MJ Research Inc.*, 2004 WL 5683983, *3 (D. Conn. Dec. 17, 2004) (dismissing antitrust case centering on large per-unit royalty that forced rival licensees to set significantly higher prices).

41. *See infra* Section III.C. (discussing the significance of reverse payments generally).

42. *See, e.g.*, *Timken Roller Bearing Co. v. United States*, 341 U.S. 593, 595 (1951) (holding global territorial division agreement was unlawful).

43. *See, e.g.*, *Gen. Talking Pictures Corp. v. W. Elec. Co.*, 304 U.S. 175, 179–80 (1938) (upholding field-of-use restraint that precluded licensees from selling competing application of patented technology). A similar restriction type limits the licensee to a particular segment of the customer base. *See, e.g.*, *United States v. Westinghouse Elec. Corp.*, 471 F. Supp. 532, 535 (N.D. Cal. 1978), *aff'd* 648 F.2d 642 (9th Cir. 1981) (upholding customer base restriction).

44. *See, e.g.*, *United States v. E.I. du Pont de Nemours & Co.*, 118 F. Supp. 41, 50 (D. Del. 1953) (noting that restricting output harms innovation and "deprive[s] society of its benefits"). For a discussion of price restraints, *see infra* notes 150–163 and accompanying text.

45. *See, e.g.*, *King Drug Co. of Florence v. SmithKline Beecham Corp.*, 791 F.3d. 388, 394–95, 397 (3d Cir. 2015) (condemning settlement in which patentee agreed not to launch generic version of its own patented drug, which would have competed with generic drug offered by rival licensee).

46. *See infra* Section III.C.

agreements, the participating firms may agree to be restrained in parallel.⁴⁷ This sometimes manifests as the joint creation of a centralized pool, which will centralize the terms on which all the firms' patents are licensed, usually by setting a royalty rate.⁴⁸

B. What's the Problem?

Why not simply prohibit any settlement that restrains competition prior to a judgment holding the patent valid and infringed? One critical reason is that frequently a settlement between rivals *must* restrain competition in order to be mutually acceptable. If litigation has a chance of excluding the patentee's competition,⁴⁹ then it has a chance of providing monopoly profits. Thus, in expectation, litigation will have an adverse effect on competition (the extent of which depends on the likelihood of validity and infringement) and will thus provide larger joint profits than open competition.⁵⁰ In any such case, a non-restrictive settlement cannot satisfy both firms; there would not be enough profits to divvy up.⁵¹

Consider a simple example in which the patentee faces a single prospective rival.⁵² Suppose that monopoly profits are \$100M, while duopoly gives joint profits of \$60M. Suppose further that the parties think the patent holder has a 50-50 chance of winning, in which case the patent holder would secure an injunction and earn the monopoly profit. Ignoring litigation costs, patent litigation leads to an expected profit level of $50\% \times \$100M + 50\% \times \$60M = \$80M$. Clearly this

47. See, e.g., *United States v. Line Material Co.*, 333 U.S. 287, 288–89 (1948) (condemning settlement in which rival patentees fixed prices at which they would require licensee-manufacturers to sell patented product).

48. *Standard Oil Co. (Ind.) v. United States*, 283 U.S. 163, 168 (1931); *Line Material Co.*, 333 U.S. at 356 (upholding pool of patents held by competitors, but only due to insufficient evidence of market power). For further discussion of cross-licensing and pooling, see *infra* Section IV.C.

49. Injunctive relief is no longer available to successful infringement claimants by default. *eBay Inc. v. MercExchange, LLC*, 547 U.S. 388, 392–93 (2006). However, when the parties are direct competitors, injunctive relief is still typically awarded. See, e.g., Christopher B. Seaman, *Permanent Injunctions in Patent Litigation After eBay: An Empirical Study*, 101 IOWA L. REV. 1949, 1990–91 (2015).

50. Throughout the analysis, it is important to distinguish between (a) the anticipated odds of hypothetical patent litigation (which is what concerns this paper); and (b) the probative weight of evidence, as appraised by a court in *actual* patent litigation. These are *ex ante* and *ex post* probabilities that cannot logically be equated. For instance, suppose the firms litigate and the court finds that the patent is 60% likely to be valid and infringed, resulting in a win for the patentee. This is very different from saying the firms believed (*ex ante*) that the patentee has a 60% likelihood of winning. If the firms knew in advance that the court would find a 60% likelihood of validity and infringement, they would impute a 100% likelihood to a win for the patentee.

51. See Erik Hovenkamp, *Competition, Inalienability, and the Economic Analysis of Patent Law*, 21 STAN. TECH. L. REV. 33, 40–41 (2018) (noting that litigation becomes rational for firms if they are limited to non-restrictive settlements); cf. Meurer, *supra* note 1, at 80.

52. The case of multiple rivals is addressed *infra* Part IV.

expected profit exceeds the \$60M that the firms would earn in a non-restrictive settlement. As a result, in order for a settlement to be mutually acceptable, it must necessarily restrain competition to some degree — namely, by enough to raise their joint earnings from \$60M to \$80M.

The example also hints at the antitrust problem, however. In order to reach a satisfactory agreement, the firms need only restrain competition enough to raise their joint profits by \$20M. Any such settlement would be proportional: it elicits the same level of joint profits as the expected result of litigation.⁵³ But from the firms' perspective, why stop there? They would maximize their joint wellbeing — the usual objective of contracting parties — by further restraining competition to the monopoly or cartel level.⁵⁴ That leaves the market less competitive and affords lower consumer welfare than the expected result of litigation.⁵⁵ Absent antitrust intervention, this must be mutually beneficial for the firms, provided that the monopoly profit is divided in a way that they both deem acceptable. And this point is independent of the parties' expectations about litigation, which will merely affect the agreed-upon division of monopoly profits.

To summarize, the antitrust problem is a tension between two facts: (1) the firms' settlement must restrain competition to *some* degree in order to be mutually acceptable; and yet (2) the firms always prefer to restrain competition all the way to monopoly. While settlements preserving monopoly will be most problematic in cases where the patent is highly likely to be invalid or noninfringed, it is a mistake to presume that the antitrust concern arises only in these cases. Rather, such probabilities merely affect the extent of the problem — the amount by which the firms' preferred settlement (shared monopoly) deviates from the proportional one.

1. Contrast with Settlements in Other Contexts

This problem distinguishes IP settlements among rivals from those occurring in virtually all other areas of private law, where no analogous concern arises. In other contexts, settlements will naturally emulate the parties' expectations about how litigation would play out. That is, an *ex ante* settlement will delimit rights in the same way they are expected

53. In principle, proportionality could rely on a different metric for competitive effects (e.g. consumer welfare effects rather than profit effects). But, for both normative and pragmatic reasons, the profit-based definition is best. See *infra* Section II.D.

54. This does not necessarily mean that the rival is excluded. In many cases, the firms attempt to share the monopoly profit through a collusive arrangement in which they all remain operational. See, e.g., *United States v. Line Material Co.*, 333 U.S. 287, 288–89 (1948) (condemning settlement in which competitors fix prices in relevant product market).

55. With respect to the consumer welfare point, see discussion *infra* Section II.D.

to end up ex post. This happens because the parties simply cannot mutually agree on anything else. Consider an example. Two neighboring businesses, *P* and *D*, are involved in a tort dispute. *P* alleges that noisy machines operated by *D* are causing a nuisance, interfering with *P*'s own business operations.⁵⁶ The parties are uncertain as to whether *D*'s activity would be deemed tortious by a court, but they know that an injunction would issue if it were construed as such. Further, suppose that *P* values the cessation of the noise at \$10k, whereas *D* would incur a loss of just \$6k by abandoning its use of the noisy machine. Therefore, the parties' joint welfare is maximized by *D* surrendering its right to use the noisy machine. And they could reap a trade surplus of \$4k by striking an agreement to that effect.

Assuming the parties can contract effectively, they will therefore enter into a settlement in which *P* pays *D* to stop using the noisy machine. This results in the same allocation of rights as an injunction, which was not certain to issue. Nevertheless, this settlement lines up exactly with the allocation of rights the parties expect to end up with following litigation: even if the court did not enjoin *D* on final judgment, *P* would simply pay *D* (some amount between \$6k and \$10k) to give up its right to use the machine ex post. And, of course, if the court did issue an injunction, the parties would not contract around this, since this judgment already provides the jointly efficient allocation of rights. Thus, the court's judgment will not have any effect on how the relevant rights are ultimately allocated ex post.⁵⁷

Thus, accounting for potential ex post contracting, there is usually no meaningful sense in which private settlements might deviate from the expected result of litigation. It makes no difference that the settlement might include a payment to secure this result.⁵⁸ Intuitively, the court's disposition will merely establish a concrete status quo — a certain allocation of rights that the parties are then free to rearrange through contract. If the settlement effects an outcome that seems unlikely to accrue from a judgment, the parties are merely anticipating the contract they expect to enter into ex post. And since that ex post agreement would be perfectly lawful, there is no cause for concern if the same result is arranged ex ante.⁵⁹

56. For a similar, well-known nuisance dispute, see *Sturges v. Bridgman* [1879] 11 Ch D 852 (Eng.).

57. This, of course, is just the Coase theorem, which says that, if transaction costs are sufficiently low to permit private contracting, the initial assignment of property rights (which could be a judgment) will not affect how such rights are ultimately allocated through the market. Ronald H. Coase, *The Problem of Social Cost*, 3 J.L. & ECON. 1, 15–16 (1960).

58. The payment is necessary in the settlement due to the fact that an injunction is not certain to issue. For instance, if *P* has a 50-50 chance of winning in tort litigation, then there is a 50% chance that *P* will have to pay *D* at least \$6k to halt its use of the machine. Thus, in their settlement, *P* will have to pay *D* at least $50\% \times \$6k = \$3k$.

59. See Hovenkamp, *supra* note 51, at 36.

By contrast, in IP settlements between competitors, the parties are subject to an atypical constraint on private bargaining: not all judgments can be lawfully contracted around after they issue. If a court deems the patent invalid or noninfringed, then the parties cannot lawfully agree that the defendant-rival will be excluded anyway (say, in exchange for cash).⁶⁰ The patent has been ruled out as a legitimate basis for such exclusion.⁶¹ More generally, when a judgment upholds a defendant-rival's right to compete freely, it implicitly bars all agreements to the contrary.

A corollary is that the Coase theorem, as conventionally stated, is ill-suited to the present context⁶² — not because there is something wrong with it, but because it implicitly assumes that all relevant rights in dispute are entirely alienable (i.e. lawfully transactable),⁶³ regardless of how they are initially allocated. This assumption fails in the present context, since a judgment may leave the defendant with an inalienable right to compete.⁶⁴ Consequently, the firms can secure the monopoly profit with certainty only through an ex ante settlement, while the defendant's right to use the invention remains presumptively alienable.

C. The Scope of the Patent Doctrine

Legal disputes over restrictive patent arrangements (including vertical ones) predate federal antitrust laws,⁶⁵ which were first codified in the Sherman Act of 1890.⁶⁶ Antitrust policy on the subject weaved a turbulent path over the subsequent hundred years.⁶⁷ However, there is one policy rubric that courts have espoused with some regularity throughout most of antitrust's history, at least until recently: the “scope of the patent” test. The test contemplates a patent's term and claim boundaries as forming a zone of commerce within which the restraint or exclusion of competition is privileged without qualification.⁶⁸

60. *Id.*

61. When there is ostensibly no potential justification for a horizontal restraint, it is deemed “naked,” and a per se rule applies. *See id.* at 48.

62. *Id.* at 40.

63. Most familiar property rights are entirely alienable. But some rights are inalienable and cannot be given up or restrained through private contracting. *See* Guido Calabresi & A. Douglas Melamed, *Property Rules, Liability Rules, and Inalienability: One View of the Cathedral*, 85 HARV. L. REV. 1089, 1092 (1972).

64. Hovenkamp, *supra* note 51, at 37.

65. *See, e.g., Adams v. Burke*, 84 U.S. 453, 456–57 (1873) (applying first-sale doctrine to condemn post-sale restraint on use of patented article).

66. 15 U.S.C. §§ 1–7 (2012 & Supp. III 2016).

67. *See* Kaplow, *supra* note 3, at 1815. *See generally* Richard Gilbert & Carl Shapiro, *Antitrust Issues in the Licensing of Intellectual Property: The Nine No-No's Meet the Nineties*, 28 BROOKINGS PAPERS ON ECON. ACTIVITY 283 (1997) (describing history of antitrust enforcement with respect to intellectual property rights).

68. *See Mercoid Corp. v. Mid-Continent Inv. Co.*, 320 U.S. 661, 665 (1944) (“[A patent] carries, of course, a right to be free from competition in the practice of the invention. But the

The scope of the patent test has faced criticism from both conservative and liberal commentators, although not always for the same reasons. One general problem is that the test stylizes itself as a collection of bright-line rules, corresponding to the boundaries on the exclusionary entitlement, yet nobody can agree on where the bright lines are.⁶⁹ In the context of horizontal settlements, the test has faced widespread criticism for providing overbroad immunity to highly restrictive arrangements deemed to fall within the scope of the patent.⁷⁰ Under the test, a settlement that prohibits the rival from using the patented invention is necessarily lawful, so long as the restriction terminates before the patent expires. It does not matter that the patent may be invalid with high probability, because the patent's validity is not an "antitrust question."⁷¹

This has led many commentators to attack the test for essentially assuming that the relevant patent is valid, notwithstanding that the settlement's terms may have been chosen based on the expectation that it probably isn't.⁷² In practical effect, the test deems a settlement to be immune from antitrust attack if it goes no farther than the most restrictive possible outcome of patent litigation, which would presumably be a permanent injunction.⁷³ Since this is a statement about mere possibility, not probability, it follows that the test is agnostic as to the likelihood that the patentee is properly entitled to an injunction.

It is worth noting, however, that in almost all other areas of private litigation, the logic underpinning the test makes perfect sense. As explained earlier, in most contexts, there is no cause for concern when a settlement essentially replicates an injunction that is not certain to issue. This makes it unnecessary to fuss over the odds that the plaintiff would

limits of the patent are narrowly and strictly confined to the precise terms of the grant."); *Coupe v. Royer*, 155 U.S. 565, 576 (1895) ("[T]he courts have no right to enlarge a patent beyond the scope of its claim . . . [and] the patentee . . . is bound by it. . . . He can claim nothing beyond it.") (quoting *Keystone Bridge Co. v. Phoenix Iron Co.*, 95 U.S. 274, 278 (1877)).

69. See, e.g., Kaplow, *supra* note 3, at 1848 ("[T]he courts seem to assume that there exists some transcendent notion of what constitutes 'normal' or 'proper' patent exploitation . . . [but] in reality, courts lack any such uniform conception of the appropriate scope of a patent.").

70. See, e.g., Michael A. Carrier, *Why the "Scope of the Patent" Test Cannot Solve the Drug Patent Settlement Problem*, 16 STAN. TECH. L. REV. 1, 1 (2012).

71. *FTC v. Actavis Inc.*, 570 U.S. 136, 169 (2013) (Roberts, C.J., dissenting) (stating that whether patent is valid "is a question of patent law, not antitrust law"). Significantly, the design-focused approach advocated here is not vulnerable to this objection, since it strips away the patent issues from the antitrust claim.

72. See Carrier, *supra* note 70, at 6 ("The problem with courts that rely on the scope test today is that they unwittingly assume that the patent is valid."); Edlin et al., *supra* note 18, at 591 (stating that scope of the patent test "effectively presumes that [the patent holder] would have won its case with certainty").

73. Accordingly, the test does arrive at the right conclusions where the firms' settlement is more restrictive than any possible judgment. See, e.g., *United States v. Line Material Co.*, 333 U.S. 287, 314–15 (1948) (condemning settlement under which competing patent holders fix product prices).

have prevailed.⁷⁴ But that is only because there is usually no possible judgment that the parties are obliged to stick with.

Thus, the scope of the patent test has some intuitive appeal for the very same reason that it is fundamentally incapable of addressing the settlement problem: it treats probability as irrelevant, because in most contexts, it really is irrelevant. But patent settlements do not comport with our usual intuitions. Here, it is not enough to frame the relevant legal constraints around the *possibility* of an exclusionary judgment. Rather, it becomes necessary to consider that the parties' settlement may restrain competition by far more than patent law would likely authorize, and that it may do so by design.

D. Actavis and Proportional Effects

In a 2003 article, Carl Shapiro was the first to formalize the idea that antitrust should require a settlement's competitive effects to be no more restrictive than the expected result of patent litigation.⁷⁵ This standard respects patent law as the ultimate authority over the balance between static competition and the rate of innovation.⁷⁶ It is patent law, not antitrust, that is tasked with formulating the conditions under which a patentee is entitled to exclude a competitor. Consistent with this, the proportional-effects rule simply makes antitrust a steward of patent law. Indeed, expectations about a hypothetical patent case are shaped entirely by patent law, as memorialized in judicial opinions, statutory language, and so on. In this way, the antitrust limits on horizontal patent settlements are not "home-grown," but rather inherited from patent law itself.⁷⁷

Another justification for the proportional-effects rule is that litigation is logically the relevant "but-for" yardstick for evaluating settlement agreements.⁷⁸ To that end, the proportional-effects rule balances the interests of patentees and consumers: a patentee can still earn the

74. See *supra* Section II.B.1.

75. Shapiro, *supra* note 1, at 391. Compare *id.* (arguing that antitrust intervention becomes appropriate when settlement provides lower consumer welfare than litigation would provide in expected value) with Meurer, *supra* note 1, at 78 (considering earlier formulation of antitrust enforcement standard that hinged on total profits reaped through settlement, independent of expected result of litigation).

76. Shapiro, *supra* note 1, at 393.

77. Of course, there are situations in which the Patent Act does not affirmatively authorize a particular practice, yet antitrust intervention nevertheless seems inappropriate. But almost all such cases involve vertical restraints, which are usually lawful even in lieu of a patent. See, e.g., Mark A. Lemley, *The Economic Irrationality of the Patent Misuse Doctrine*, 78 CALIF. L. REV. 1599, 1605 (1990); Gilbert & Shapiro, *supra* note 67, at 287–89. By contrast, in the context of horizontal settlements, the operative restraints would be unlawful (generally *per se*) but for the patent.

78. *In re Cipro Cases I & II*, 348 P.3d 845, 863–64 (Cal. 2015) ("What would the state of competition have been without the agreement? . . . [T]he relevant comparison is with the average level of competition that would have obtained . . . if the parties had litigated . . .").

profit it anticipates getting through litigation (in addition to avoiding the costly and lengthy litigation process), whereas consumers are no worse off than expected if the patent were litigated to judgment.⁷⁹ Finally, the proportional-effects rule helps to ensure that rewards are commensurate with inventive contributions, which is universally recognized as an important objective of the patent system.⁸⁰ If a given patent is likely invalid — say, because it is obvious in light of prior art — then it is fitting that the patentee’s reward should be smaller than if the invention were not obvious at all. The proportional-effects rule achieves this by ensuring that patents enjoy a “less durable” monopoly when they are more likely to be invalid.⁸¹

In Shapiro’s original proposal, a proportional settlement is one that generates the same amount of consumer welfare as the expected result of litigation.⁸² By contrast, Lemus and I propose a metric that hinges instead on total profits.⁸³ They are similar because a reduction in competition will be reflected by both an increase in total profits and a reduction in consumer welfare. These two effects are thus alternative metrics with which one can measure anticompetitive impact. The two metrics coincide for some kinds of settlements.⁸⁴ Where they differ, the profit-based conception is actually modestly better for consumers.⁸⁵ In these cases, a proportional settlement still gives the same profits as the expected result of litigation, but it yields slightly greater consumer welfare than litigation would provide in expectation.⁸⁶ This is one of the benefits of the profit-based formulation.⁸⁷ A second benefit is that this formulation is inherently better positioned to evaluate how settlement design affects private bargaining. That is because profit considerations

79. See Shapiro, *supra* note 1, at 396.

80. See, e.g., Michael B. Abramowicz & John F. Duffy, *The Inducement Standard of Patentability*, 120 YALE L.J. 1590, 1597 (2011); Ayres & Klemperer, *supra* note 2, at 1008–09; Kaplow, *supra* note 3, at 1817–18.

81. Erik Hovenkamp, *Challenge Restraints and the Scope of the Patent*, 1 ANTITRUST CHRONICLE 48, 53 (2016) (“This result — that lower quality patents are less durable and thus impose smaller restraints in commerce — performs a socially valuable function. Judgments on patent validity are binary . . . [b]ut patent quality is [not] [But validity disputes have the capacity to] ensure that commercial restraints are somewhat proportionate with patent quality.”).

82. Shapiro, *supra* note 1, at 391 (proposing that “a settlement must leave consumers at least as well off as they would have been from ongoing patent litigation”).

83. Hovenkamp & Lemus, *supra* note 7, at 7–8.

84. In particular, they coincide for restraints on entry (e.g. delayed entry). *Id.* at 24–26.

85. *Id.* This is driven by a technical feature of Marshallian demand, namely that consumer surplus is convex (in price) whereas producer surplus and total surplus are concave.

86. A royalty is an example of a settlement where the proportional outcomes leave consumers strictly better off than the expected result of litigation. See *id.* at 9–10.

87. *Id.* at 7–8.

are ultimately what drive the firms' negotiations. Put differently, to know what the firms can agree on, we must focus on profits.⁸⁸

1. Pay-for-Delay Settlements

The proportional-effects rule caught on in large part because it is a very intuitive approach for evaluating a particular kind of settlement format that gained widespread antitrust attention beginning in the early 2000s: "pay-for-delay" settlements, also known as "reverse payment" settlements. These agreements had become very prevalent in the pharmaceutical industry⁸⁹ and had been widely accused of unlawfully restraining market entry by generic drug makers.⁹⁰ A comprehensive understanding of the pay-for-delay problem requires consideration of the pharmaceutical regulatory environment, which controls important aspects of the process by which generic firms challenge brand-name drug patents, usually on validity grounds.⁹¹ These regulatory issues have been discussed extensively in the literature.⁹² For present purposes, a short summary will be sufficient; a later section discusses what this paper's analysis suggests about the current regulatory regime.⁹³

When a brand-name drug is covered by a patent, a generic firm can attempt to enter the market early by challenging the patent as invalid.⁹⁴ To maintain incentives to challenge weak drug patents, the Hatch-Waxman Act offers an inducement to generic challengers. The first generic drug maker to file for FDA approval (the "first-filer") will be entitled to 180 days of "generic exclusivity" in the event that it successfully

88. Notwithstanding these points, the consumer welfare versus profits distinction probably will not matter much in practice. They are likely too similar for real-world courts to be able to distinguish between them.

89. It is not clear whether such settlements have been executed in other industries. This is in part because mandatory FDA filings create a paper trail that makes it easy to identify settlements that delay generic entry, whereas in other industries there is no such paper trail. For further discussion, see generally Erik Hovenkamp & Jorge Lemus, *Delayed Entry Settlements at the Patent Office*, 54 INT'L REV. L. & ECON. 30 (2018).

90. See, e.g., FTC, PAY-FOR-DELAY: HOW DRUG COMPANY PAY-OFFS COST CONSUMERS BILLIONS 1 (2010) (finding that pay-for-delay agreements "delay generic competition" and "consumers lose [because] they miss out on generic prices that can be as much as 90 percent less than brand prices"). For a broader discussion of antitrust issues in the pharmaceutical industry, see Arti K. Rai, *Fostering Cumulative Innovation in the Biopharmaceutical Industry: The Role of Patents and Antitrust*, 16 BERKELEY TECH. L. REV. 813, 814 (2001).

91. Drug Price Competition and Patent Term Restoration (Hatch-Waxman) Act, Pub. L. No. 98-417, 98 Stat. 1585 (1984). The most important provisions appear in 21 U.S.C.A. § 355 (West 2018).

92. See, e.g., Hemphill, *supra* note 18, at 101 (explaining how certain aspects of the Hatch-Waxman Act contribute to the pay-for-delay problem); Michael A. Carrier, *Solving the Drug Settlement Problem: The Legislative Approach*, 41 RUTGERS L.J. 83, 83–85 (2009); Crane, *supra* note 18, at 757–59.

93. See *infra* Section IV.A.

94. This involves the generic firm making a "Paragraph IV Certification," which is an assertion that the brand-name drug is patented, but that the patents are either invalid or not infringed by the generic drug. 21 U.S.C.A. § 355(j)(2)(A)(vii)(IV).

invalidates the patent.⁹⁵ During this time, no “later-filing” generics can enter the market.⁹⁶ While the brand-name firm (the patent holder) operates as a monopolist, it can set very high prices and earn substantial profits. But if the generic’s challenge proves successful, there will be very intense price competition, particularly after the generic exclusivity period concludes, since the generic versions of the drug are therapeutically equivalent to the brand-name drug. As a consequence, prices and profits will fall substantially. Thus, the firms can earn much larger profits if they preserve the monopoly and share in the proceeds.

To accomplish this, the brand-name firm agrees to give a “reverse payment” (conventionally a cash lump sum) to the generic firm.⁹⁷ In exchange, the latter agrees to terminate its challenge and delay its entry into the market for some number of years, often until soon before the patent expires. For example, suppose the brand-name drug maker, *B*, would earn \$100M over the patent term if it retains its monopoly. If the generic firm, *G*, enters the market and competes freely, total profits would fall to \$40M over that period, with each of the firms earning \$20M. Suppose the firms think the patent has a 50-50 chance of being valid. Then the expected result of litigation involves profit levels of $50\% \times \$100M + 50\% \times \$20M = \$60M$ for *B*, and $50\% \times \$20M = \$10M$ for *G*. However, rather than litigating, the firms agree that *G* will stay out of the market until the patent expires, and *B* will make a reverse payment of \$25M. This is a highly disproportionate outcome that leaves both firms with far greater earnings than they expected to get through litigation.

As a means of proving that the settlement’s effects were more restrictive than the expected result of litigation, many authors advocate an inference-based approach. It relies on the size of the reverse payment (in conjunction with a long delay period) as a proxy for the likelihood

95. *Id.* § 355(j)(5)(B)(iv).

96. A major problem, however, is that third-party generics (i.e. those other than the first-filer) lose their incentive to challenge the patent, for even if such challenge proves successful, it is still the first-filer who will get the exclusivity reward. For a richer discussion, see Hemphill, *supra* note 18, at 108.

97. The name “reverse payment” reflects that, in conventional patent settlements, any stipulated payments flow in the opposite direction.

that the patent is invalid.⁹⁸ This has been dubbed the “Actavis Inference.”⁹⁹ The logic underpinning this inference is simply that if the firms really thought the patentee were likely to win, then it would not have to make such a large payment to persuade the rival to accept a lengthy delay period.¹⁰⁰ Note, however, that this would be a moot point under the scope of the patent test, since pay-for-delay settlements are no more restrictive than a permanent injunction.¹⁰¹

2. The *Actavis* Decision

The antitrust debate over pay-for-delay reached its climax in the Supreme Court’s 2013 *Actavis* decision. Over the objections of three dissenters, the majority held that pay-for-delay settlements can indeed violate the antitrust laws.¹⁰² The majority embraced the logic behind the proportional-effects rule,¹⁰³ describing the relevant violation as an effort “to prevent the *risk* of competition.”¹⁰⁴ Further, consistent with what many commentators had urged, the Court held that a large, unexplained reverse payment “can provide a workable surrogate for the patent’s weakness, all without forcing a court to conduct a detailed exploration of the . . . patent itself.”¹⁰⁵

In another passage, the majority noted that the firms are perfectly free to settle in other ways and expressly mentioned a payment-free

98. See, e.g., Shapiro, *supra* note 1, at 408 (“[A] naked [reverse] payment . . . (in excess of avoided litigation costs) is a clear signal that the settlement is likely to be anticompetitive.”); Edlin et al., *supra* note 18, at 610 (“[T]ypically a settlement with a large reverse payment will also injure consumers compared with [the expected result of litigation].”); Carrier, *supra* note 18, at 75 (“[S]trong evidence of invalidity is presented by payments from brands to generics that exceed what the generics could have gained by entering the market.”). But see Gregory Dolin, *Reverse Settlements as Invalidity Signals*, 24 HARV. J.L. & TECH. 281, 282–84 (2011) (arguing that the inference properly warrants reexamination of the patent, rather than antitrust intervention); Joshua B. Fischman, *The Circular Logic of Actavis*, 66 AM. U. L. REV. 91, 91 (2016) (arguing it is inappropriate to draw inference of invalidity from firms’ subjective beliefs about likelihood of invalidity).

99. Edlin et al., *supra* note 18, at 585.

100. *Id.* at 591–92.

101. See Keith Hylton, *Antitrust and Intellectual Property: A Brief Introduction*, in THE CAMBRIDGE HANDBOOK OF ANTITRUST, INTELLECTUAL PROPERTY, AND HIGH TECH 81, 90 (Roger Blair & Daniel Sokol, eds., 2017) (“The typical reverse payment settlement . . . would be lawful under the scope-of-the-patent test.”). Prior to *Actavis*, a number of lower court decisions relied on the scope of the patent test to conclude that pay-for-delay agreements are immune to antitrust attack. See, e.g., *FTC v. Watson Pharm., Inc.*, 677 F.3d 1298, 1312 (11th Cir. 2012).

102. See *FTC v. Actavis Inc.*, 570 U.S. 136, 140 (2013).

103. *Id.* at 154–57 (noting that “the relevant anticompetitive harm” is the elimination of “the risk of competition,” and that antitrust intervention is appropriate where “the patentee seeks to induce [using a payment] the generic challenger to abandon its claim with a share of its monopoly profits that would otherwise be lost in the competitive market”).

104. *Id.* at 157 (emphasis added).

105. *Id.* at 158.

delayed-entry settlement as one such option.¹⁰⁶ This alternative arrangement is the “pure-delay” settlement discussed in the introduction, which generally elicits proportional effects.¹⁰⁷ Although the Court did not note this point, that it identified pure delay as a lawful option suggests that it was mindful that settlement design is indeed germane to the antitrust analysis.

Elsewhere in the opinion, the majority rejected the dissent’s contention that antitrust’s purview is properly limited to agreements that fail the scope of the patent test.¹⁰⁸ However, the Court declined to apply a per se rule to reverse payments, as the FTC had urged.¹⁰⁹ Instead, the Court concluded the payments must be evaluated under the rule of reason.¹¹⁰

E. The Administrability Problem

The proportional-effects rule merely specifies a threshold for antitrust enforcement. How to administer the rule in practice is a separate question entirely. To that end, intuition suggests it would be very difficult to administer a standard that depends on expectations about a patent judgment that never issued.¹¹¹ The Actavis Inference — reliance on a large reverse payment as a proxy for the patent’s weakness — is a viable workaround. But this requires that the settlement agreement supplies a suitable proxy. As such, the viability of the Actavis Inference is seemingly limited to cases involving reverse payments. This is a major limitation, given that many patent settlements that have faced antitrust scrutiny over the years did not involve such payments.¹¹²

Moreover, as discussed in a later section,¹¹³ even after *Actavis*, some courts remain highly skeptical of this proxy-based justification

106. *Id.*

107. For a discussion of the regulatory issues commonly associated with such settlements, see *infra* Section IV.A.

108. *Actavis*, 570 U.S. at 147 (noting that even though reverse payment settlement agreement might fall “within the scope of the exclusionary potential of the [defendant’s] patent,” this does not “immunize the agreement from antitrust attack”); see also *In re Cipro Cases I & II*, 348 P.3d 845, 856 (Cal. 2015) (noting scope of the patent test “accords excess weight to the policies motivating patent law, gives insufficient consideration to the concerns animating antitrust law, and must be rejected”).

109. 570 U.S. at 159–60.

110. *Id.* (noting that likelihood of anticompetitive effects depends on size of payment relative to litigation costs and plausibility of any legitimate explanations asserted by defendant).

111. See Dolin, *supra* note 98, at 317 (noting many scholars have highlighted apparent administrability challenges).

112. See, e.g., *Gen. Talking Pictures Corp. v. W. Elec. Co.*, 304 U.S. 175 (1938); *United States v. Gen. Elec. Co.*, 272 U.S. 476 (1926); *Ethyl Corp. v. Hercules Powder Co.*, 232 F. Supp. 453 (D. Del. 1963); *United States v. Singer Mfg. Co.*, 374 U.S. 174 (1963). These cases involved a wide range of different restraints and other provisions, but no reverse payments.

113. See *infra* Section V.B.

for avoiding direct scrutiny of the patent.¹¹⁴ This is likely to remain problematic, for the evidentiary weight of the relevant inference is somewhat vulnerable to the subjective appraisals of judges. For instance, judges who are skeptical of the Actavis Inference may require extremely large payments in order to find a violation or may make it relatively easy for defendants to rebut the inference. For instance, one court recently held that the bald suggestion by amici that the patentee might be risk-averse was sufficient to rebut the Actavis Inference.¹¹⁵

When there is no suitable proxy, enforcement would seem to require a direct assessment of the patent issues within the antitrust adjudication. While such an assessment may seem sensible in theory, it would be deeply problematic in practice.¹¹⁶ Assessment of patent issues requires a court to formulate a prediction about the odds of a hypothetical litigation (which would center on issues of pure patent law) to ascertain what the settlement's competitive effects should be compared against. In light of the difficulties this would present, several authors¹¹⁷ and courts¹¹⁸ have taken to calling this the "turducken" approach.¹¹⁹ This is just a particular embodiment of the more general problem of attempting to decide "a case within a case," which is an awkward and difficult process that courts tend to regard as unpalatable.¹²⁰ Both patent litigation and antitrust litigation are highly complex on their own terms. To force both undertakings into a single adjudication would substantially chill enforcement efforts. Moreover, the court's determinations on the patent issues would ostensibly constitute fact-finding, making it more difficult to rectify erroneous determinations on appeal.

114. See, e.g., *In re Wellbutrin XL Antitrust Litigation Indirect Purchaser Class*, 868 F.3d 132, 168 (3d Cir. 2017) ("While the size of a reverse payment may have some relevance . . . it is far from dispositive.").

115. See *id.* Incidentally, the Supreme Court had expressly rejected this argument in *Actavis*, 570 U.S. 136, 172 (2013).

116. See Shapiro, *supra* note 1, at 397 ("[O]ne key practical problem" is that this approach "requires an informed judgment as to the strength of the patent(s) at issue.").

117. See Edlin et al., *supra* note 18, at 619 (arguing that in pay-for-delay cases the "turducken" problem can be largely avoided since reverse payment provides suitable proxy for answering antitrust question).

118. *United Food & Commercial Workers Local 1776 & Participating Emp'rs Health Welfare Fund v. Teikoku Pharma USA*, 296 F. Supp. 3d 1142, 1155 (N.D. Cal. 2017) ("I disagree that plaintiffs need to prove in this case that [defendant] would have won its patent litigations. That turducken is not only unappetizing as a matter of judicial efficiency, it is not required (or even suggested) by the *Actavis* opinion."); *FTC v. Watson Pharm., Inc.*, 677 F.3d 1298, 1315 (11th Cir. 2012) ("[A]ttempt[ing] to decide how some other court in some other case at some other time [would] have resolved some other claim if it had been pursued to judgment . . . would be deciding a patent case within an antitrust case about the settlement of the patent case, a turducken task."), *rev'd sub nom.* *FTC v. Actavis, Inc.*, 570 U.S. 136 (2013).

119. Turducken is a duck stuffed into a chicken stuffed into a turkey (ideally post-mortem).

120. See, e.g., Polly A. Lord, *Loss of Chance in Legal Malpractice*, 61 WASH. L. REV. 1479, 1482–84 (1986); Robert M. Klein, Andrew M. Feldman, & Mark Sullivan, *Legal Malpractice Claims — Advanced Strategies for Case-Within-a-Case Litigation*, 43 OHIO N.U. L. REV. 331, 332–36 (2017).

III. THE ECONOMICS OF SETTLEMENT DESIGN: AN OVERVIEW

The administrability problem is driven by an implicit assumption that pervades the literature, which is that one cannot apply the proportional-effects rule without some kind of evidence or indicia relating to the patent's validity. This is true even in the pay-for-delay context, where it manifests as the presumption that we need a proxy to stand in for a more rigorous patent inquiry. This has given rise to an empirical approach under which there is little or no focus on precisely *how* competition is restrained, so long as the *extent* of it seems reasonably commensurate with the expected result of litigation.¹²¹

But in fact, there is a much more effective option, which simultaneously avoids the case-within-a-case problem and enables the proportional-effects rule to be administered to all types of settlements: focus on the settlement design, and how it affects private bargaining *in general* — regardless of the odds that any particular patent is valid and infringed. In our recent economics paper, Jorge Lemus and I develop a very general economic approach demonstrating that a settlement's design determines what competitive outcomes the firms can mutually agree on in relation to their expectations about how litigation would play out.¹²² This Section provides a distilled, nontechnical summary of those economic points, with emphasis on simple numerical examples and links to the case law. For simplicity, the analysis below will assume that an injunction would fully exclude the patentee's rival from the marketplace and that this would result in monopoly.

The underlying economic approach is a method of analyzing private bargaining between actual or potential competitors in the shadow of patent litigation. It can be applied to any model of oligopoly competition (which specifies the details of the firms' competitive relationship),¹²³ which is chosen at the discretion of the modeler. That means the general approach is not married to any particular modeling assumptions about the relevant competitive environment. Rather, one has the freedom to choose a model and attendant assumptions that seem most fitting in a given case.

121. See, e.g., Shapiro, *supra* note 1, at 399–400 (considering settlement that leads price level to rise above competitive level, but remaining silent as to how firm achieved this result).

122. Hovenkamp & Lemus, *supra* note 7, at 7–11.

123. See *id.* at 18–19. Note that, while most of the paper assumes that the firms earn equal profits under open competition, this is just for expositional simplicity and can be relaxed. A later section of the paper considers a more general model that allows for the firms to earn asymmetric profits even when competition is unrestrained. See *id.* at 26–29.

A. Relevant Aspects of Settlement Design

The term “settlement design” is merely meant to capture the relevant types of provisions being negotiated. Focusing on these details helps to clarify what kinds of provisions will lead to problems and which ones will not. In broad outline, the relevant aspects of settlement design are those that influence how total profits are distributed. There are three categories of provisions that do so. First, all settlements relevant to this paper include some restraint on the patent holder’s rival. This is a “licensing restraint,” since it is appended to a license granted to the rival.¹²⁴ As noted earlier, there are many different types of horizontal restraints, such as territorial restrictions, field-of-use restraints, delayed entry, etc. To that end, the first relevant aspect of a settlement is the particular manner in which the rival-licensee is restrained.¹²⁵

A settlement may (but need not) further include one of two other types of provisions. The first is a reverse payment. This is not necessarily cash; it may be any form of valuable consideration that does not directly affect competition, but which does so indirectly by skewing the bargain over the licensing restraint. For example, pay-for-delay is a settlement format involving delayed entry as the relevant licensing restraint, and it further includes a reverse payment. By contrast, “pure delay” is an alternative design that includes the same licensing restraint, but no reverse payment.

Second, the settlement may include a counter-restraint: a separate restraint on the patentee’s right to commercialize its own invention. Unlike the licensing restraint, in this case the restrained activity is already definitively lawful by default. For example, the patentee might surrender its right to sell a particular version of the patented good,¹²⁶ or agree not to operate in a particular territory. As explained below, reverse payments and counter-restraints are two sides of the same coin; they both destroy proportionality by providing absolute flexibility in the sharing (i.e. redistribution) of total profits.

124. See, e.g., DEP’T OF JUSTICE & FED. TRADE COMM’N, ANTITRUST GUIDELINES FOR THE LICENSING OF INTELLECTUAL PROPERTY 7 (2017) (discussing various licensing restraints).

125. Aside from such restraint, a rival may also agree to pay a lump sum license fee upfront. See, e.g., Alain Bosquet et al., *Risk Sharing in Licensing*, 16 INT’L J. INDUS. ORG. 535, 536 (1998).

126. For instance, as an alternative to reverse payments, some pharmaceutical patent holders have offered generic rivals a promise not to launch an “authorized generic” of the brand-name drug, which would otherwise compete with the rival’s generic offering. See, e.g., Edlin et al., *supra* note 18, at 597–98 (discussing no-authorized-generic provisions). For other examples, see *infra* Section III.C.

B. Licensing Restraints

This section addresses the relevance of the particular manner in which the rival is restrained. In this analysis, we will assume the settlement does not include a reverse payment nor any counter-restraint on the patentee. Recall that all licensing restraints have the same two overarching effects on competition: (1) total profits rise, relative to the competitive level; and (2) the rival earns a smaller percentage of total profits, because the restraint leaves it at a competitive disadvantage. For instance, as a royalty rate increases, the rival is left at an increasingly large cost disadvantage. Similarly, if the rival's operations are limited to certain territories while the patentee continues to operate everywhere, this too reflects a disadvantage for the rival.

Why does it matter *how* the rival is restrained? One might conjecture that restraints differ in terms of the extent to which they can diminish competition. Perhaps some are capable of producing a total monopoly, while others are not capable of going quite that far. But, in fact, this is not the sense in which restraints differ from one another. Every restraint can elicit any feasible level of joint profits — that is, any level between the competitive and monopoly bookends — by adjusting the *magnitude* with which the restraint is applied. For every type of restraint, there is some relevant variable (the magnitude) that governs the intensity of its anticompetitive effects and, by extension, the extents to which total profits rise and consumer welfare falls. For example, if the restraint is a royalty, the magnitude is simply the dollar-value of the royalty rate; under a delayed-entry restraint, the magnitude is the length of the delay period; and if the restraint is an output limitation, the magnitude is the particular cap imposed on the rival's sales.¹²⁷

Once the firms have decided what kind of licensing restraint they will use, they bargain over the magnitude. Every restraint preserves the competitive equilibrium when its magnitude is negligibly small. But the market converges to monopoly as the magnitude grows ever-higher. For example, a delay period that persists until patent expiration results in monopoly over the full remainder of the patent term, but a delay period of twenty minutes has no discernable impact on competition. Similarly, a royalty rate of a few pennies will have negligible competitive effects, but as it increases, it eventually becomes so large that the rival cannot afford to set a price that any consumers would willingly pay: this too results in monopoly.

If any restraint can be used to elicit any feasible level of joint profits, how do they differ from one another? The answer is that they differ

127. As this illustrates, magnitude is measured in various kinds of units, depending on the nature of the restraint. This is why it is convenient to focus on a generalized notion of magnitude. The important point is just that competition declines and total profits rise as the magnitude grows larger.

only with respect to how joint profits are *distributed* between the firms as the magnitude increases. As such, in the context of horizontal settlements, one can conceptualize a licensing restraint as nothing more than a specification of how (each feasible level of) total profits would be distributed between the firms.¹²⁸ This, in turn, determines what the firms can mutually agree on. To illustrate, suppose that the monopoly profit is 100, and that each firm earns a profit of 25 under open competition. Thus, any restraint can generate total profits of any amount between 50 and 100. But different restraints can distribute such levels differently. For example, consider two settlement arrangements that employ alternative restraints, where both settlements are used to generate a total profit of 70. Thus, the two alternative restraints are being applied with equivalent magnitudes. But they may nevertheless allocate that total profit of 70 differently, such that each firm earns different profits in each settlement. It may then be that one of the settlements is mutually-agreeable, while the other is not.

To illustrate, consider two hypothetical licensing restraints, *A* and *B*, which are simply characterized by alternative profit distributions, as described in the table below.¹²⁹ The column for each restraint corresponds to a settlement format that relies on that restraint alone. This will allow us to compare these alternative restraints on proportionality grounds based on how they impact private bargaining.

Table 1: Hypothetical Licensing Restraints

	Total Profits	Profit Distribution (patentee, rival)	
		Restraint <i>A</i>	Restraint <i>B</i>
Magnitude ↓	50	(25, 25)	(25, 25)
	60	(40, 20)	(37, 23)
	70	(55, 15)	(50, 20)
	80	(70, 10)	(64, 16)
	90	(85, 5)	(79, 11)
	100	(100, 0)	(100, 0)

128. Edlin et al., *supra* note 18, at 598 (showing that restraints can be reduced to specification of how each level of total profits would be distributed between firms).

129. For simplicity, the table considers profit increases in increments of ten. But more generally, the choice of restraint determines how all attainable levels of total profits would be distributed between the firms.

As we move down the table, each restraint is being applied more forcefully and hence its overall anticompetitive effects are increasing. This is why total profits are rising (and, although not displayed in the table, consumer welfare is falling). I have given the restraints generic names because for now, the point is simply to show that it is the restraints' distributional characteristics that drive bargaining outcomes and proportionality. The next subsection explains how we apply this analysis to evaluate specific kinds of restraints, which effectively boils down to determining what distribution numbers go in each restraint's profit tables.

To evaluate the restraints on proportionality grounds, we consider different possible litigation odds, and in each case we ask what the parties could mutually agree on. We then compare these bargaining possibilities to the expected result of litigation. For instance, suppose that the patent is 40% likely to be held valid and infringed. Then expected total profits from litigation are $40\% \times 100 + 60\% \times 50 = 70$. Thus, a proportional settlement outcome is one that elicits total profits of 70. To discern what the parties can agree on, we must compute the individual (per firm) profits expected to accrue from litigation. They are $40\% \times 100 + 60\% \times 25 = 55$ for the patent holder and $60\% \times 25 = 15$ for the rival. To be mutually acceptable, a proportional settlement would have to provide the same allocation of profits. After all, there is no alternative way to distribute a total profit of 70, while still ensuring that the patentee gets at least 55 and the rival gets at least 15.

Can the parties agree on the proportional outcome? Consider restraint *A* first. By inspection of Table 1, the proportional outcome (the row with total profits of 70) is indeed mutually acceptable under restraint *A*. This settlement outcome provides the same allocation of profits that is expected to accrue from litigation. But might the firms nevertheless agree to restrain competition disproportionately? The answer is clearly no. As restraint *A* is applied with greater magnitude, the rival's profits fall. If the magnitude were increased beyond the proportional level, the rival's profits would fall below 15 — an unacceptable result. On the other hand, the firms could never agree on any settlement that provides a joint profit lower than 70, since any mutually agreeable outcome must provide at least that amount.

Thus, under restraint *A*, the firms *can only agree on the proportional outcome*. If we repeat this analysis for different possible litigation odds, we always arrive at the same reassuring conclusion.¹³⁰ For instance, suppose the patentee instead has a 60% chance of winning.

130. One can confirm this point for any litigation odds that generate one of the total profit levels listed in the table. As Lemus and I show, there are numerous restraints that always induce proportional outcomes, just like restraint *A*. See Hovenkamp & Lemus, *supra* note 7, at 21–24.

Then expected litigation payouts become 70 for the patentee and 10 for the rival. Then under restraint *A* the firms can only agree on the outcome involving a total profit of 80, which is the proportional outcome now that the patentee is 60% likely to win.

Thus, restraint *A* keeps the firms honest, no matter what odds they happen to impute to litigation. It follows that, so long as the firms use such a restraint, we have no cause to worry about a disproportionate result; the firms could not mutually agree on any such terms. This is ultimately because the restraint always distributes total profits in the same way that litigation does in expectation. As a consequence, a proportional restraint cannot diminish competition excessively without becoming unacceptable to the rival.

What about restraint *B*? Notably, in comparison to *A*, restraint *B* is always more generous to the rival, and, by extension, stingier to the patentee. That is, the distributions under *B* are skewed in the rival's favor (relative to *A*). This will alter what the firms can mutually agree on for any litigation odds. Suppose that, as before, the patent is 40% likely to be valid and infringed, so that the proportional outcome involves a total profit of 70. The firms cannot agree on this outcome; the patentee would earn less than the profit of 55 that it anticipates getting through litigation. Thus, in order for the firms to agree, they must restrain competition disproportionately. Under restraint *B*, they would ratchet up the restraint's intensity until total profits reach 80, which exceeds the competitive level by 50% more than would the proportional result. Even if we change the litigation odds, the firms will continue to restrain competition beyond the proportional level, for they are unable to agree on anything else.

This illustrates another important point: "excessive" restraints compel the firms to mutually agree on a disproportionate result, because they cannot mutually agree on anything less.¹³¹ The reason is very straightforward. In order for the rival to agree to be restrained excessively, it must receive more favorable distributions of profits. But that necessarily results in the patentee obtaining *less* favorable profit allocations. Thus, at the proportional magnitude, the patentee is not getting enough to match its expected payout from litigation. The result is that bargaining possibilities are distorted toward higher magnitudes, and hence private bargaining must always lead to disproportionate anticompetitive effects.

Figure 1 below shows how bargaining possibilities compare to the proportional outcome for all possible litigation odds under restraints like *B*. The x-axis variable is the probability that the patent is valid and infringed (denoted P_{vi}). The y-axis gives total profits, which capture

131. See *id.* at 18–24 (analyzing these "excessive" restraints and showing that they necessarily prevent the firms from mutually agreeing on the proportional outcome).

the restraint magnitude. The joint profits under open competition are denoted Π^C , while the monopoly profit is denoted Π^M . For any particular value of P_{vi} , imagine a vertical line drawn upward from P_{vi} 's position on the x-axis. It intersects the diagonal black line at the proportional level of total profits. But this is not mutually agreeable. Instead, the mutually agreeable outcomes are those at which the vertical line crosses the checkered strip. As the graph shows, bargaining possibilities always involve disproportionate outcomes (reflected by excessive total profits, which imply excessive magnitudes). By contrast, a proportional restraint (like restraint *A*) would generate bargaining possibilities that coincide exactly with the diagonal black line, reflecting that the firms cannot mutually agree on anything other than the proportional outcome.

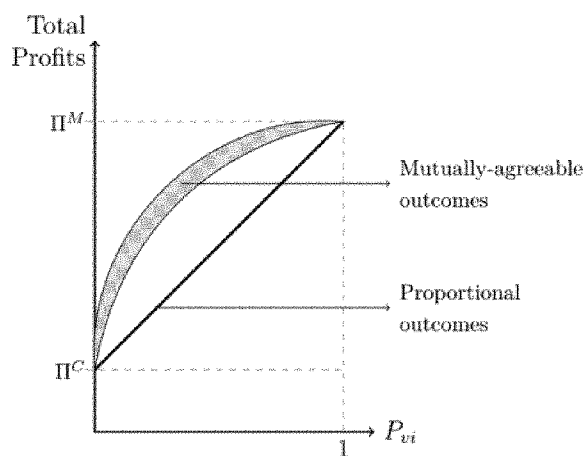


Figure 1: Bargaining Possibilities Under an Excessive Restraint

A final point helps to illustrate that all restraints that firms might actually implement will end up influencing bargaining either: (1) just like restraint *A*; or else (2) much like *B* (although not necessarily identically). Restraint *B* differs from *A* in that its distributions are skewed in the rival's favor. The only possible alternative would involve distributions that instead skew in the patentee's favor (relative to *A*). But, in fact, in that case the firms could not mutually agree on *any* outcomes, and hence no settlement could rely on such a restraint. It is easy to see why this happens. Under restraint *B*, the rival is willing to be restrained beyond the proportional level. But if the distributions were skewed toward the patentee, the opposite result occurs: the rival is not willing to be restrained even to the proportional level; it demands a lesser magnitude. But any mutually acceptable settlement must generate at least the

proportional level of total profits in order to match both firms' expected litigation payouts. Thus, no such restraint could support a mutual agreement.

The implication is that any viable restraint will either be proportional, just like *A*, or disproportional, approximately like *B*. The latter qualification arises because problematic restraints like *B* can differ in the particular extent to which they distort bargaining possibilities away from the proportional outcome, depending on how strongly they skew the profit distributions away from the proportional ones. But, for purposes of establishing liability under the proportional-effects rule, the relevant question is simply whether or not the restraint forces a disproportionate result.¹³²

1. Applications

Restraints *A* and *B* are just hypothetical examples. The point was to illustrate how, all else being the same, alternative restraints lead to different bargaining outcomes by virtue of distributing profits differently. Of course, in practice, we want to inquire into specific, real-world restraints — an output restriction, for example. To do this, one must model the output restraint in order to determine how it distributes profits in equilibrium. In effect, this is what tells us what numbers would go in the profit table for the output restraint. One can then repeat this process for other types of restraints in order to compare them on proportionality grounds, just as we did with *A* and *B*.

For instance, consider a licensing restraint that limits the rival's entry in some way, such as by delaying it or by limiting it to some territories. (Recall that we are currently assuming there is no reverse payment or counter-restraint.) Under conventional market conditions, these restraints are proportional, just like restraint *A* above.¹³³ To illustrate, consider the pure delay settlement format, wherein the rival's entry is delayed. As noted in the introduction, this restraint always induces proportional outcomes. This point has already been recognized by a number of antitrust commentators in the literature on pay-for-delay settlements, where these alternative formats are often contrasted.¹³⁴ The caveat, discussion of which is left to a later section,¹³⁵ is that the Hatch-Waxman Act undermines the proportionality of pure delay settlements

132. Of course, if the distortion is quite small, it may not be possible to prove in court. But this is not a unique problem; for all kinds of anticompetitive conduct, if the relevant harm is relatively small, it may be impossible to prove. This just reflects the practical difficulties of litigation generally.

133. See Hovenkamp & Lemus, *supra* note 7, at 16–17 (explaining that under ordinary market conditions, entry restraints always elicit proportional outcomes and always provide same consumer welfare as expected result of litigation).

134. See, e.g., Shapiro, *supra* note 1, at 407–08; Hemphill, *supra* note 18, at 1588–90.

135. See *infra* Section IV.A.

reached in the pharmaceutical industry,¹³⁶ which is the commercial environment that entry settlements are best known for.¹³⁷ But as explained later, this is fundamentally a Hatch-Waxman problem, not a problem with the entry restraint itself, which should be regarded as an appropriate and lawful option.¹³⁸

Consider the pure delay format in a context that is not plagued by Hatch-Waxman's problematic regulatory measures. Suppose that there are ten years remaining in a patent term, that monopoly generates a profit of 10 per year, and that open competition gives each firm an annual profit of 3. Thus, an injunction would result in total monopoly profits of 100 over the patent term, while invalidation of the patent would accrue a competitive joint profit of 60 (30 per firm) over that span.

Under a pure delay settlement, the firms simply bargain over a delay period. The monopolist's overall settlement payout — that is, the total profits it gets over the remainder of the patent term — is given by (1) the monopoly annual profit for each year of the delay period, plus (2) the competitive annual profit for each subsequent year in the patent term. The rival's overall settlement payoff is just the amount corresponding to part (2), as it earns nothing during the delay period. This formula tells us what level of total profits and corresponding distribution would accrue from any possible delay period the firms might choose. The table below displays these profit numbers, although, for simplicity, it restricts focus to delay periods that are multiples of two years.¹³⁹

136. See Elhaage & Krueger, *supra* note 18, at 314 (showing that in pharmaceutical markets, Hatch-Waxman Act leads pure delay to elicit disproportionate effects on competition); Hemphill, *supra* note 18, at 1590 (noting that model with conventional market conditions “fits pharmaceutical regulation poorly”); Hovenkamp & Lemus, *supra* note 89, at 33 (explaining that Hatch-Waxman regulations lead otherwise-proportional restraints to diminish competition excessively).

137. Other kinds of entry settlements — in particular territorial restrictions and field-of-use restraints — are common in many other industries, however. See, e.g., *Sec. Materials Co. v. Mixermobile Co.*, 72 F. Supp. 450, 454 (S.D. Cal. 1947) (upholding legality of territorial restraint); *Gen. Talking Pictures Corp. v. W. Elec. Co.*, 304 U.S. 175, 180 (1938) (holding no antitrust violation by field-of-use restriction limiting what commercial uses licensees can engage in), *aff'd on reh'g*, 305 U.S. 124 (1938).

138. Rather than condemning entry restraints, policymakers should focus on resolving the problems that currently pervade the pharmaceutical regulatory environment. For discussion of potential fixes for these regulations, see generally Carrier, *supra* note 92; Crane, *supra* note 1; Mark A. Lemley & Scott C. Hemphill, *Earning Exclusivity: Generic Drug Incentives and the Hatch-Waxman Act*, 77 ANTITRUST L.J. 947 (2011).

139. Note that the table's profit measures are cumulative over the ten-year period remaining in the patent term.

Table 2: Pure Delay Settlements with a Single Rival

Delay Period (magnitude)	Total Profits	Profit Distribution (patentee, rival)
0	60	(30, 30)
2	68	(44, 24)
4	76	(58, 18)
6	84	(72, 12)
8	92	(86, 6)
10	100	(100, 0)

As before, the restraint's magnitude (now explicitly given by the delay period) is increasing as we move down the table. We can thus utilize the table to analyze bargaining and proportionality just as we did above. Suppose the parties think the patent has a 40% likelihood of being valid and infringed. Then, in expected value, litigation would generate a total profit of $40\% \times 100 + 60\% \times 60 = 76$ over the balance of the patent term. Thus, by inspection of Table 2, a proportional settlement would involve a four-year delay period. It is easy to verify that, just like restraint *A* from the preceding section, this proportional settlement gives each firm exactly the same profit it expected to get from litigation.¹⁴⁰ Thus, it is mutually acceptable and, following the same logic as before, it is the only outcome the firms can agree on. Here too, we reach the same conclusion under a different specification of the possible litigation odds.¹⁴¹ It is also easy to see that pure delay also leads to the same consumer welfare as the expected result of litigation.¹⁴²

Other kinds of entry restraints, such as territorial limitations or field-of-use restrictions, will also be proportional in this way.¹⁴³ That is because they affect the distribution of profits in substantially the same way, notwithstanding that they limit entry in different dimensions. For instance, consider a territorial restraint in which the rival is constrained

140. Expected profits are $40\% \times 100 + 60\% \times 30 = 58$ for the patent holder, and $60\% \times 30 = 18$ for the rival.

141. For example, if the patentee is 60% likely to win in court, expected total profits are 84, with 72 going to the patentee and 12 to the rival. This induces the firms to agree on a six-year settlement, which gives the same total profit and is therefore proportional.

142. As with total profits, the ensuing consumer welfare is just a probability-weighted average of the consumer welfare levels corresponding to monopoly and open competition. Other authors have also noted this point. *See, e.g.,* Hemphill, *supra* note 18, at 1588–89 (suggesting “settlement splitting entry in accordance with the probabilities [of validity and infringement] has the same effect upon expected patentee profits, [rival] profits, and consumer welfare” as the expected result of litigation).

143. *See* Hovenkamp & Lemus, *supra* note 7, at 24.

to operate in a particular subset of U.S. states, within which the rival can compete freely.¹⁴⁴ Suppose these states comprise 40% of the overall domestic consumer base. The patentee, by contrast, operates everywhere¹⁴⁵ and therefore enjoys a monopoly in 60% of the country.¹⁴⁶ Thus, the market is divided 60-40 between monopoly and duopoly. But that is exactly like the four-year delay settlement from the table above. They both divide the market between monopoly and duopoly, but one does so territorially while the other does so temporally.¹⁴⁷

Royalties are also proportional under conventional market conditions, although this is harder to illustrate without a formal model.¹⁴⁸ This suggests that a royalty-based licensing restraint — provided it is not accompanied by any other concerning provisions — does not warrant antitrust scrutiny, even if it happens to have a strong effect on competition (which happens when the royalty rate is sufficiently high). Strong anticompetitive effects would only occur if the parties think the patent is likely to be valid and infringed.¹⁴⁹ By contrast, if the settlement relies on price or output restraints,¹⁵⁰ it will necessarily lead to disproportionate bargaining outcomes.¹⁵¹ In this case, the results are just like those we found under restraint *B*, above.¹⁵²

144. The Patent Act authorizes a patentee to grant a license “to the whole or any specified part of the United States.” 35 U.S.C. § 261 (2012).

145. This point is critical. If the patentee agrees to stay out of those territories in which the rival operates, this would be a disabling restraint on the patentee. And the result of the agreement would be no different than that of naked market division, which is ordinarily illegal per se. *See, e.g., Palmer v. BRG of Georgia, Inc.*, 498 U.S. 46, 49–50 (1990) (holding nonpatent territorial division agreement, in which each firm operated exclusively in its respective territory, was illegal per se).

146. This requires the assumption that the settlement can effectively prevent competition between the competitive territories and the monopoly ones.

147. *See Hovenkamp & Lemus, supra* note 7, at 23–24.

148. *Id.* at 9–10 (showing that royalties always elicit proportional bargaining outcomes in standard Cournot model of oligopoly competition).

149. For instance, in *Applera* the court rightly dismissed an antitrust claim centering on a large per-unit royalty charged by one firm to many of its rivals. *Applera Corp. v. MJ Research Inc.*, No. 3:98cv1201 (JBA), 2004 WL 5683983, at *12 (D. Conn. Dec. 17, 2004).

150. The courts have consistently held that a patentee has the right to limit the sales of a rival licensee. *See, e.g., Atari Games Corp. v. Nintendo of Am., Inc.*, 897 F.2d 1572, 1578 (Fed. Cir. 1990) (holding output restraint on rival-licensee is not anticompetitive as matter of law); *Ethyl Corp. v. Hercules Powder Co.*, 232 F. Supp. 453, 460 (D. Del. 1963) (stating patentee can “control the commercial production of the unpatented product by the legitimate manipulation of his patent monopoly”). A price restraint imposed on a rival licensee is technically lawful under limited circumstances. *United States v. Gen. Elec. Co.*, 272 U.S. 476, 489–90 (1926). *But see FTC v. Actavis, Inc.*, 570 U.S. 136, 150 (2013) (suggesting *General Electric* does not reach any settlements other than those involving single patentee and single licensee). However, if the settlement involves several competing patentees restraining their prices in parallel, it is unlawful. *See, e.g., United States v. Line Material Co.*, 333 U.S. 287, 291–93, 314–15 (1948).

151. Hovenkamp & Lemus, *supra* note 7, at 18–24.

152. For an explanation of why an output cap leads competition to be restrained excessively, see *id.* at 11–14. In effect, both a royalty and an output cap reduce the challenger’s output. In the royalty case, the rival is made to pay a “tax” to the patentee, but there is no

In fact, price and output restraints can also lead to a much more troubling situation, in which the rival may actually *benefit* from being restrained to some degree.¹⁵³ This is like a worse version of restraint *B*, wherein the distribution is so skewed that the rival's profits are initially going up as the magnitude increases. Unlike restraint *B*, the rival may actually earn greater profits when it is restrained to some degree than when it is not restrained at all — albeit not too much, since every restraint will eventually diminish the rival's profits to zero when the magnitude gets sufficiently high. To illustrate, consider a price restraint applied to the rival's product alone, which specifies what price the rival-licensee must apply to its sales of the patented good.¹⁵⁴ Note that, in an ordinary price-fixing cartel, the participating firms agree to raise their prices in parallel, and this results in greater profits. The difference with the present price restraint is that it only applies to one of the firms, namely, the rival; this means that the patent holder can undercut the rival, leaving the latter with fewer sales.

However, whether or not this will leave the rival with lower profits than it earns under open competition depends on: (1) the extent of the stipulated price increase; and (2) the aggressiveness of inter-party price competition, which depends on market conditions (such as the extent to which the firms' products are differentiated).¹⁵⁵ If price competition is not too intense and the restraint does not prescribe too large a price increase, then the patentee will not set its own price too much lower than the rival's. Moreover, if the rival-licensee retains the right to challenge the patent (which it does by default),¹⁵⁶ and if the patent is likely invalid, then the patentee has no incentive to aggressively undercut the rival's price. To do so would induce the rival to challenge the patent. This is likely to result in invalidation, which would preclude any restraints on competition. Thus, despite the fact that only the rival's price is restrained, the deal may be profitable for both firms. In such cases, it resembles a sort of asymmetric price-fixing in which the rival's price may be somewhat higher (and its market share somewhat smaller) than the patentee's, but not so much as to prevent both firms from earning supracompetitive returns.

A case from the early 20th century provides an illustration of this phenomenon. In *United States v. General Electric*,¹⁵⁷ the Supreme

payment requirement under an output cap. Thus, the distribution of profits is always more favorable to the rival under an output cap than under a royalty.

153. *Id.* at 28 (analyzing “accommodating” restraints under which a rival-licensee can earn larger profits when it is restrained than when it is not).

154. More generally, the restraint could stipulate a floor on what price the rival can set.

155. Hovenkamp & Lemus, *supra* note 7, at 23 n.45 (demonstrating that price restraints can enhance rival-licensee's profits under price competition with differentiated products).

156. *See Lear, Inc. v. Adkins*, 395 U.S. 653, 670 (1969) (holding licensee is not automatically estopped from challenging validity of licensed patent).

157. 272 U.S. 476 (1926).

Court confronted an agreement between patentee General Electric and its rival-licensee, Westinghouse.¹⁵⁸ General Electric held several patents pertaining to lightbulbs with tungsten filaments, and it was a dominant manufacturer in the lightbulb market at the time. Under the agreement, Westinghouse was licensed to sell the patented bulbs subject to the requirements that Westinghouse must: (1) sell its bulbs at a price stipulated by General Electric,¹⁵⁹ and (2) pay only a small royalty (2% of revenues) provided that its market share remained no greater than 15%.¹⁶⁰

The Supreme Court upheld the agreement as lawful, concluding that the right to stipulate a rival-licensee's price falls within the scope of the patent.¹⁶¹ Over the subsequent hundred years, the decision has been the subject of widespread criticism,¹⁶² and there is reason to doubt that it would survive renewed scrutiny before the Supreme Court.¹⁶³ In any case, for present purposes, the point is that Westinghouse may well have earned larger profits while being price-restrained than it would have obtained by competing freely.

C. Reverse Payments and Counter-Restraints

The preceding section focused on settlements that include only a restraint on the rival-licensee. In these cases, holding all else constant, the restraint alone determines what the firms can agree on in relation to their expectations about litigation. That is because the restraint inherently limits the possible distributions of joint profits. However, if the settlement design further includes a reverse payment or counter-restraint,¹⁶⁴ then the firms face no such constraint. Rather, either such provision provides absolute flexibility as to how total profits are divided. As a result, the choice of licensing restraint becomes irrelevant, for the settlement format's distributive characteristics can always be modified arbitrarily to suit the firms' interests.

158. *Id.* at 478.

159. *Id.* at 479 (noting agreement required Westinghouse to "follow prices and terms of sale from time to time fixed by the Electric Company").

160. As noted in the lower court opinion, the royalty rate would increase to 10% if Westinghouse's market share rose above 15%. *United States v. Gen. Elec. Co.*, 15 F.2d 715, 718 (N.D. Ohio 1925).

161. 272 U.S. at 485.

162. See, e.g., Kaplow, *supra* note 3, at 1846; cf. Gerald R. Gibbons, *Domestic Territorial Restrictions in Patent Transactions and the Antitrust Laws*, 34 GEO. WASH. L. REV. 893, 906 (1965).

163. See, e.g., *Asahi Glass Co., Ltd. v. Pentech Pharm., Inc.*, 289 F. Supp. 2d 986, 992 (N.D. Ill. 2003) ("The Court upheld the arrangement [in *General Electric*] but I doubt that it would do so today . . .").

164. Recall that these terms were defined in the discussion above accompanying Section III.A.

1. Reverse Payments

Much of the literature on reverse payments, by scholars on both sides of the policy debate, centers on what such payments allow us to infer about the likelihood that the relevant patent is invalid.¹⁶⁵ This section's analysis will not address this point. Instead, our focus is simply on how reverse payments and counter-restraints affect the bargaining process generally, regardless of what odds the firms might impute to litigation. As explained below, both such provisions utterly destroy proportionality. The fundamental reason for this is that they permit the firms to bargain separately over the size and distribution of total profits.

The intuition for the latter point is very simple. Consider a silly analogy in which the firms are bargaining over a pie, which will be cut into two parts, one for each firm. When the settlement design includes a licensing restraint alone, as in the previous section, the size and cut of the pie are inexorably linked: the firms cannot enlarge the pie without also altering the cut (in the patentee's favor),¹⁶⁶ and the rival's share shrinks to zero as the pie reaches its maximal size. Thus, the firms cannot mutually agree to maximize the size of the pie — notwithstanding that they both want as much as possible — because the rival will not accept a serving size of zero: it must get no less than it expects to obtain through litigation. But, if the firms can bargain independently over the size and cut of the pie, then of course they will always agree to make it as big as possible, regardless of their litigation expectations. No matter how they might have divided a smaller pie, they can always divide the largest one in a way that leaves them both with larger portions.¹⁶⁷ This is precisely what reverse payments and counter-restraints enable the firms to do.

The juxtaposition of pure delay and pay-for-delay provides a good illustration of this. Consider the same numerical example from the last section's discussion of pure delay.¹⁶⁸ The patent term has 10 years remaining, and monopoly profits are 10 per year, while competition results in annual profits of 3 per firm. Table 3, below, compares the profit distributions for each of several delay periods attained by pure delay

165. Compare Edlin et al., *supra* note 18, at 585 (“[A] large and otherwise unexplained payment, combined with delayed entry, supports a reasonable inference of harm to consumers . . .”) with Crane, *supra* note 1, at 702 (claiming direction of payment “is a poor proxy for determining the ex ante likelihood that the plaintiff would prevail at trial”).

166. As the restraint is applied with greater magnitude (increasing total profits), the rival is placed at a greater competitive disadvantage. This leaves the patent holder with a larger share of a larger pie. This is illustrated in the profit tables above, as the rival's profit share is always shrinking as the magnitude increases.

167. For instance, they could divide the larger pie into the same percentage-shares as the smaller one.

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

and pay-for-delay. For simplicity, the table considers three possible sizes of the reverse payment (abbreviated as “RP”).¹⁶⁹

Table 3: Profit Distribution for Pure Delay and Pay-For-Delay

Delay Period (magnitude)	Profit Distribution (patentee, rival)			
	Pure Delay (RP = 0)	Pay-for-Delay		
		RP = 10	RP = 20	RP = 30
0	(30, 30)	(20, 40)	(10, 50)	(0, 60)
2	(44, 24)	(34, 34)	(24, 44)	(14, 54)
4	(58, 18)	(48, 28)	(38, 38)	(28, 48)
6	(72, 12)	(62, 22)	(52, 32)	(42, 42)
8	(86, 6)	(76, 16)	(66, 26)	(56, 36)
10	(100, 0)	(90, 10)	(80, 20)	(70, 30)

Suppose that the patent is 40% likely to be valid and infringed, so that expected litigation profits are 58 and 18 for the patentee and its rival, respectively. In this case, the five boldface table entries are those settlement possibilities that are mutually preferred to litigation. As before, under pure delay, the proportional outcome (in this case, a four-year delay) is the only one the firms can agree on. But private bargaining will never lead to this result under pay-for-delay. For each of the three settlements involving non-zero reverse payments, the firms cannot mutually agree on the proportional outcome; rather, all outcomes that the firms mutually prefer to litigation are disproportionately restrictive. More specifically, with a reverse payment, the four-year delay period becomes unacceptable to the patentee; it now demands a longer delay period to justify the payment. Simultaneously, the rival is now willing to be restrained beyond the proportional level.

Much of this analysis is reminiscent of what we found in comparing restraints *A* and *B* in the previous section. When there is a reverse payment, the profit distributions are skewed in the rival’s favor, increasingly so as the payment grows. Similarly, we found that restraint *B* induced disproportionate bargaining outcomes because its profit distributions are skewed in the rival’s favor. However, in two highly problematic respects, the reverse payment settlement format differs from the one that included restraint *B* alone: here, (1) the rival still gets to share in total profits when competition is restrained all the way to monopoly; and (2) the firms have absolute control over the extent of such sharing and can therefore bargain separately over this variable. It thus becomes

169. There is no total profit column in Table 3, but total profits (for each delay period) are the same as in Table 2.

easy to see that whenever a reverse payment provision is added to any baseline settlement design — even one that is otherwise proportional — the result is to destroy proportionality absolutely.

To further unpack this, note that the table has two shaded cells, both involving the maximal ten-year delay period. These settlement outcomes, and no others, are Pareto efficient as between the firms. Aside from being mutually preferred to litigation, each of these outcomes has the property that there is no alternative that both firms like better (nor any alternative that one firm likes better and the other is equally happy with). One of the most fundamental concepts in economics is that private contracting leads to Pareto efficient outcomes.¹⁷⁰ By contrast, the other three boldface settlement outcomes are Pareto *inferior*. While they are all mutually preferred to litigation, in each case there is an alternative and more restrictive outcome (namely, one of the Pareto efficient ones) that would leave both firms better off. As such, absent the prospect of antitrust penalties, the firms would be irrational not to choose one of the two Pareto efficient settlement outcomes.¹⁷¹

Note that a reverse payment is not itself a “restraint.” Rather, it is a tool for distorting the bargain over the licensing restraint. It “distorts” the bargain by altering what magnitudes the firms can mutually agree on. With a reverse payment, the mutually-agreeable outcomes shift, becoming more restrictive (and more profitable for both firms) as the payment grows larger. By contrast, if the relevant side payment were a lump sum license fee — paid by the rival-licensee to the patent holder — precisely the opposite would happen: the mutually-agreeable bargaining outcomes shift toward less restrictive outcomes.¹⁷²

This clarifies the error in claims that the direction of a side-payment is not very relevant to the antitrust analysis.¹⁷³ To be fair, however, such claims were raised in the context of evaluating what a payment allows us to infer about the “true” litigation odds in a particular case. But the above discussion shows that the payment direction’s

170. This proposition, which holds under very general conditions, is the First Fundamental Welfare Theorem of economics. See, e.g., Mark Blaug, *The Fundamental Theorems of Modern Welfare Economics, Historically Contemplated*, 39 HIST. POL. ECON. 185, 185 (2007); see also Edlin et al., *supra* note 18, at 614 (“The distinction between *feasible* outcomes and actual *equilibrium* outcomes is fundamental to neoclassical economic analysis . . .”) (emphasis added). In the present context, an equilibrium is a settlement outcome that is (1) mutually-preferred to litigation; and (2) Pareto efficient among the possible outcomes that are attainable under the operative settlement design.

171. Which of these two outcomes the firms will agree on is indeterminate, since the firms have opposing preferences over them. But, for antitrust purposes, it does not matter which one they choose, since both are maximally restrictive.

172. Hovenkamp & Lemus, *supra* note 7, at 27. Note however that this does not imply the firms can just rely on lump sum license fees to reach a mutual agreement. Such agreement requires that competition be restrained far enough to match the firms’ litigation expectations, whereas lump sum license fees do not restrain competition at all. See *supra* note 36.

173. See, e.g., Crane, *supra* note 1, at 700–02 (arguing that antitrust literature overstates significance of direction in which payment flows).

impact on private bargaining is hardly ambiguous and has nothing to do with the litigation odds. In particular, the payment's direction determines whether it distorts the mutually-agreeable outcomes toward stronger or weaker anticompetitive effects. Nothing could be more relevant to the antitrust analysis.

2. Counter-Restraints

When a settlement includes a counter-restraint, the result is that both firms are restrained (although not necessarily in the same way or to the same extent). Since that is how ordinary cartels function, the antitrust concerns are perhaps more obvious than those created by reverse payments. The courts have confronted a wide range of counter-restraints over the years. For instance, in *United States v. National Lead*,¹⁷⁴ several competitors entered into a licensing arrangement under which the firms divided territories world-wide, with each firm designated as the exclusive seller within a given set of territories.¹⁷⁵ In *United States v. Masonite Corp.*,¹⁷⁶ several competitors entered into a patent settlement under which they all agreed to fix the prices of their products.¹⁷⁷ To similar effect, in *In re Summit Technology, Inc.*¹⁷⁸ two competing producers of laser eye surgery equipment agreed that they would both pay large royalties to a jointly-founded licensing pool (forcing their prices upward in parallel), and then share in the pool's earnings.¹⁷⁹

One such settlement arrangement has gained widespread attention in the aftermath of the *Actavis* decision. As an alternative to cash payments, some pharmaceutical patent holders (brand-name drug sellers) have opted to offer prospective rivals (generic drug makers) something else — namely, a promise not to launch their own “authorized generic” version of the drug,¹⁸⁰ which would otherwise compete with the rival-launched generic.¹⁸¹ This can be viewed as an output counter-restraint

174. 332 U.S. 319 (1947).

175. *Id.* at 319 (condemning licensing agreement that divided global territories among firms). The Patent Act expressly authorizes the grant of exclusive licenses within “the whole or any specified part of the United States.” 35 U.S.C. § 261 (2012). For discussion of agreements with domestic territorial restraints, see generally Gibbons, *supra* note 162.

176. 316 U.S. 265 (1942).

177. *Id.* at 265–66.

178. 127 F.T.C. 208 (1999).

179. *Id.* at 210. This FTC action was eventually settled, with the firms agreeing to dissolve their collusive arrangement.

180. An authorized generic is one sold by the brand-name patentee, who has already obtained FDA approval to sell the relevant drug. Importantly, an authorized generic can be launched during the first-filing generic's 180-day exclusivity period. For further discussion, see generally FTC, AUTHORIZED GENERICS: AN INTERIM REPORT (2009).

181. See, e.g., *King Drug Co. of Florence v. SmithKline Beecham Corp. (King Drug)*, 791 F.3d 388, 394 (3d Cir. 2015) (noting settlement involving “no-authorized generic” promise is unlawful under *Actavis* decision).

on the patentee, albeit one that is limited to a generic version of the patented drug.¹⁸²

Counter-restraints lead to substantially the same results as reverse payment.¹⁸³ The primary difference is that reverse payments facilitate monopoly, while counter-restraints facilitate cartels. But the important point is that both kinds of provisions permit the firms to bargain separately over the size and distribution of total profits. Recall that when the settlement restrains only the patentee's rival, this skews the distribution of profits in the patentee's favor. This asymmetry grows as the licensing restraint is applied with greater magnitude. But when both firms are restrained, this distributional asymmetry is diminished or eliminated, depending on the relative magnitudes with which the two restraints are applied. In effect, when the patentee shoulders part of the settlement's "restrictive burden," the resulting distribution of profits becomes more egalitarian.¹⁸⁴ Thus, by properly adjusting the two restraints, the firms can divide any level of the total profits however they like.

A simple illustration involves territorial market division.¹⁸⁵ The firms' settlement will divide the country between the two firms, with each firm operating exclusively in its respective territory. Suppose that monopoly profits are 100. Then if they split the country in half, they each get profits of 50; if they divide the country 60-40, then one firm gets a profit of 60 while the other gets 40, and so on. In all cases, the full monopoly profit is preserved. But, by choosing how the country will be divided, the firms can divide this profit however they like. Accordingly, this arrangement is not meaningfully different from a reverse payment settlement that delays entry until the end of the patent term. The patentee's own territorial confinement (the counter-restraint), which leaves the rival with a partial monopoly, is the analogue to the reverse payment. And if the counter-restraint is removed, the settlement format becomes proportional, just like pure delay.¹⁸⁶

182. In principle, the patent holder could further agree to limit the sales of its brand-name variant. This would enable the firms to further restrain competition and thereby share in an even larger profit. However, the antitrust concerns would be particularly salient in this case, which is perhaps why the relevant patentees have limited the restraint to sales of authorized generics.

183. As such, some courts and commenters even refer to counter-restraints (specifically, no-authorized-generic pledges) as "reverse payments." See, e.g., *King Drug*, 791 F.3d at 394 (noting no-authorized-generic commitment "is a payment to eliminate the risk of competition"). For discussion of such settlements, see generally Michael A. Carrier, *Eight Reasons Why "No-Authorized-Generic" Promises Constitute Payment*, 67 RUTGERS UNIV. L. REV. 697 (2015).

184. That is, the new division of profits is more balanced than if a licensing restraint alone had been used to generate the same level of total profits.

185. Recall that the Patent Act permits the granting of territorially-restricted licenses. See *supra* note 8.

186. See *supra* notes 146–151 and accompanying text.

3. The Pareto Problem

The problem posed by reverse payments and counter-restraints is very general: under *any* settlement design that includes either such provision, *all* Pareto efficient bargaining outcomes restrain competition to monopoly, no matter the firms' litigation expectations.¹⁸⁷ Absent the prospect of antitrust penalties, it would be contrary to both firms' interests to settle for anything less than a full monopoly profit, regardless of what they think about the patent.¹⁸⁸ It makes no difference what kind of licensing restraint is imposed on the rival.¹⁸⁹

In effect, this is the worst possible property that a settlement design can exhibit. All such arrangements are "maximally disproportional." They ensure that the firms' interests become completely aligned as to the settlement's overall restrictiveness. By contrast, in order for settlements to be proportional (or even approximately so), it is absolutely critical that the firms maintain adverse interests as to how restrictive the settlement will be.¹⁹⁰ That is, we want the patentee's rival to strive to be restrained as little as possible — not simply to fight for a larger share of the monopoly profits. Whenever a settlement design fails to maintain this adversity, we can immediately rule out the possibility that private bargaining will elicit proportional effects.

A corollary is that, if antitrust simply caps the permissible size of reverse payments — such as by requiring that they be no larger than the cost of litigation¹⁹¹ — then settling firms will simply restrain competition to the most disproportional extent that the cap permits them to

187. Hovenkamp & Lemus, *supra* note 7, at 28.

188. This is a simple point that many other authors have observed in the pay-for-delay context. See, e.g., Edlin et al., *supra* note 18, at 615 ("[I]f all reverse payments are allowed, then the equilibrium settlement will delay competition until the patent expires."). However, these prior works couch the antitrust analysis in terms of the inference that reverse payments afford, whereas the present article focuses on how the inclusion of reverse payments undermines the bargaining process in general in all cases and for any odds of validity.

189. For example, in the case of reverse payments, if the licensing restraint had been a cap on the rival's sales, the only Pareto-efficient outcomes would involve the patentee making a reverse payment so large that the rival will agree to make no sales at all (i.e. a cap of zero). This is no different from a delay period that persists until patent expiration.

190. Such adversity is usually maintained in settlements that include only a licensing restraint. See, e.g., *infra* Table 1 (showing that rival will always fight to be restrained as little as possible (under both restraints *A* and *B*), since its profits fall as the restraint is applied with greater magnitude).

191. Some articles suggest this as a relevant cap. See, e.g., Edlin et al., *supra* note 18, at 590. However, this is merely a conservative benchmark used to support an inference that competition is restrained beyond the expected result of litigation. The cap allows for such an inference even if the rival has all of the bargaining power — an extreme possibility that would allow the rival to extract the patentee's litigation cost without being delayed excessively. However, under an approach that focuses on how reverse payments affect bargaining generally, there is much less reason for the cap, absent some indication that patent challengers tend to maintain far more bargaining power than patentees.

agree on. This would mitigate only the extent to which firms' settlements are disproportional (at least in cases involving weak patents).¹⁹² This is easy to see in Table 3. Suppose that the law imposes a cap of 10 on reverse payments. Assuming as before that the patent is 40% likely to be valid, a six-year delay period (with $RP = 10$) becomes the new (and only) Pareto efficient settlement outcome among the lawful options. By contrast, without any cap, both Pareto efficient outcomes (the shaded cells) delay entry for the full remainder of the patent term. As such, limiting antitrust intervention to cases involving "large and unjustified" reverse payments does not actually resolve the antitrust problem; it just makes the problem less pronounced.¹⁹³

Similarly, any conjecture that reverse payments might serve some efficient purpose, so long as they are not too large, must be rejected as a potential justification for such dealings.¹⁹⁴ The firms will not limit the reverse payment to the supposedly-efficient level; any contrary suggestion is a fantasy. The firms are profit-maximizers, and they will behave as such. This critique is discussed in greater detail in a later section.¹⁹⁵

IV. MULTIPLE ENTRY, HATCH-WAXMAN, AND OTHER CONSIDERATIONS

The foregoing analysis has focused on the case in which the patentee faces potential competition from a single rival. But in most contexts, it is not realistic to suppose that just one rival would enter the market.¹⁹⁶ Fortunately, the antitrust analysis extends to the case of multiple entry with minimal added difficulty.¹⁹⁷ This section briefly outlines the significance of multiple entry, with the caveat that the instant discussion will set aside the regulatory issues arising under Hatch-Waxman, the effects of which are left to a later section.¹⁹⁸ We will also return focus to settlements that do not include reverse payments or

192. If the patent is relatively strong (so that, for instance, an eight-year delay period would be proportional), then a comparatively small payment could still be sufficient to generate the maximal (ten-year) delay period.

193. Cf. Michael A. Carrier, *Why a "Large and Unjustified" Payment Threshold is Not Consistent with Actavis*, 91 WASH. L. REV. 109, 110 (2016) (arguing that *Actavis* decision does not make "large and unjustified" payment necessary element of antitrust claim, although it may be relevant factor).

194. For instance, some commentators suggest that some reverse payments should be permitted on the theory that the patentee might be risk-averse. See, e.g., Marc G. Schildkraut, *Patent-Splitting Settlements and the Reverse Payment Fallacy*, 71 ANTITRUST L.J. 1033, 1061 (2004).

195. See *infra* Section V.C.

196. See, e.g., Edlin et al., *supra* note 18, at 588–94 (discussing significance of multiple entry in context of pay-for-delay settlements, demonstrating that *Actavis* Inference remains valid under multiple entry).

197. Hovenkamp & Lemus, *supra* note 7, at 28–29.

198. See *infra* Section IV.A.

counter-restraints, as any settlements including these provisions continue to undermine bargaining in the same way identified above.¹⁹⁹

Under the prospect of multiple entry, three significant differences emerge. First, if there are many prospective rivals, then invalidation may result in substantial competition and low profits, which can undermine incentives to challenge the patent in the first place.²⁰⁰ Second, there may be a “coordination problem” that makes it difficult or impossible for the patentee to settle with multiple rivals.²⁰¹ However, these two considerations are not, strictly speaking, antitrust problems. They stem from frictions that may prevent the firms from being able to settle in the first place.²⁰² But such frictions are institutional, not products of firm conduct.²⁰³

The third effect of multiple entry does bear on the antitrust analysis, however. Suppose that multiple rivals have an incentive to challenge the patent, so that their litigation threats are credible. Then the patentee must settle with all of them.²⁰⁴ The result would be that all of the rivals are restrained in parallel. The problem is that this may make each rival less uncomfortable being restrained. Intuitively, if not for the fact that the patentee is not being restrained, this arrangement might be no different from an ordinary cartel, which restrains all firms in parallel and thereby benefits all of them.²⁰⁵ By the same token, each rival benefits from the fact that its “co-rivals” are similarly restrained. The result is ultimately that excessive restraints, such as price or output restraints,²⁰⁶ elicit even more disproportionate effects in the case of multiple entry.²⁰⁷

Fortunately, proportional restraints continue to operate as such under multiple entry.²⁰⁸ This is easy to illustrate by returning to the example of pure delay (Table 2) and extending it to allow for several rivals. As before, suppose there are 10 years remaining in the patent term and that monopoly profits are 10 per year. But now suppose the patentee faces potential competition with three rivals — r_1 , r_2 , and r_3 . Finally,

199. Hovenkamp & Lemus, *supra* note 7, at 27–29 (demonstrating that reverse payments and counter-restraints destroy proportionality; such proofs extend seamlessly to the case of multiple entry).

200. Litigation costs can create the same problem. *See infra* Section IV.B.

201. For instance, each rival may reasonably be concerned about whether other settling rivals are getting better licensing terms. *See infra* notes 209–210 and accompanying text.

202. Of course, they are plainly germane to competition policy writ large. But, if the complications of multiple entry prevent any settlements from occurring in the first place, there is no potential basis for antitrust liability.

203. This point is relevant to the discussion of both Hatch-Waxman and litigation costs. *See infra* Sections IV.A. and IV.B.

204. Note that this involves separate agreements with each rival.

205. This is easiest to see when the relevant restraint applies to price or output, since an ordinary cartel operates by either limiting output or fixing high prices.

206. *See supra* notes 150–163 and accompanying text.

207. Hovenkamp & Lemus, *supra* note 7, at 29.

208. *Id.*

suppose that open competition generates a total profit of 4 per year (1 per firm). The table below gives the possible profit distributions over a range of possible delay period.

Table 4: Pure Delay Settlements with Three Rivals

Delay Period (years)	Total Profits	Profit Distribution (patentee, r_1, r_2, r_3)
0	40	(10, 10, 10, 10)
2	52	(28, 8, 8, 8)
4	64	(46, 6, 6, 6)
6	76	(64, 4, 4, 4)
8	88	(82, 2, 2, 2)
10	100	(100, 0, 0, 0)

Suppose that the patent is 40% likely to be valid and infringed. Then expected litigation profits are 46 for the patentee and 6 for each of the three rivals. Just as before, this corresponds to a four-year delay period, which is the only thing the firms can agree on. As a result, so long as the firms use a proportional restraint, a multiplicity of settlements will not create a risk that competition will be restrained disproportionately.

In fact, this has important implications for the possible coordination problem mentioned above. It may be difficult for the patentee to settle independently with each of several rivals.²⁰⁹ But, so long as the firms use a proportional settlement format, there may be little risk in letting the rivals coordinate in the negotiations, perhaps by permitting them to appoint a representative to bargain on behalf of all of them.²¹⁰ After all, under a proportional settlement, we do not have to worry that they might have an interest in being restrained excessively. For instance, in the pure delay example above, the rivals would obviously not

209. Each rival cannot fully evaluate how profitable its settlement will be without knowing what terms are being offered to third-party rivals. That could make it harder to secure rivals' agreement. Additionally, if the number of rivals is large, the transaction costs would add up quickly.

210. Of course, it may be appropriate (and perhaps advisable) for an antitrust authority to maintain some supervision over this process. An alternative to literal coordination would be to permit the challengers to rely on most-favored nation (MFN) clauses such that the patentee promises each rival that, if better terms are subsequently offered to someone else, then they must be offered to said rival as well. This way, the rivals end up receiving common terms, despite not coordinating with each other directly. However, MFN clauses can raise antitrust concerns in their own right. See Jonathan B. Baker & Judith A. Chevalier, *The Competitive Consequences of Most-Favored-Nations Provisions*, 27 ANTITRUST 20, 20 (2013).

derive any value from staying out of the market six years rather than four. Rather, absent a reverse payment, they will stay out no longer than four years. This is precisely what a single rival would demand under the same circumstances. In either case, the ultimate result is proportional.

A. Clarifying the Hatch-Waxman Problem

As noted earlier, the Hatch-Waxman Act necessarily plays a prominent role in the analysis of pay-for-delay settlements.²¹¹ The Act attempts to use generic exclusivity to incentivize generics to challenge brand-name drug patents, but the inducement mechanism is badly conceived. It actually reduces the incentive to challenge the brand-name drug patent for most generic firms — namely, all but the first-filer.²¹² That is because, even if the patent is invalidated by a later-filing generic, the inducement still goes to the first-filer.²¹³ As such, for later-filers, a patent challenge is actually less appealing than it would be but for the Hatch-Waxman Act.²¹⁴

The result is that, even if a sizeable number of generics would ultimately enter upon the patent's invalidation (albeit after the first-filer's exclusivity period), the patent holder may be able to settle with just a few of them.²¹⁵ The first-filer, along with later-filing generics that have already made investments in obtaining FDA approval, are the only ones with credible litigation threats.²¹⁶ And, because the Hatch-Waxman Act always gives the challenge inducement to the first-filer, there are fewer later-filing generics who have an interest in making such investments. Thus, in many instances, fewer generics will enter the market during the balance of the patent term than would have entered if the patent were invalidated.

This undermines private bargaining, leading otherwise proportional settlement formats to restrain competition excessively. For example, as several authors have observed, even pure delay may lead to disproportionate results as a result of the Hatch-Waxman Act's faulty

211. See *supra* Section II.D.1.

212. See Hemphill, *supra* note 18, at 1555–62.

213. *Id.* at 1583 (“Even if the first-filer loses, withdraws, or settles, a subsequent-filer does not become eligible for the bounty.”); Hovenkamp & Lemus, *supra* note 89, at 33 (“If there were no statutory exclusivity . . . a later-filer would at least be able to enter immediately if it prevailed in challenging the patent. But this is not so under [Hatch-Waxman].”).

214. *Id.*

215. In many pay-for-delay cases, the patent holder settled not only with the first-filer, but also with a few others. See Edlin et al., *supra* note 18, at 603–06. Also, there can be multiple first-filers if two or more generics all file for FDA approval on the same day. Hemphill, *supra* note 18, at 1583.

216. Hemphill, *supra* note 18, at 1582.

inducement system.²¹⁷ But this is not a problem with pure delay specifically. On the contrary, *every* settlement format will enable more restrictive bargaining outcomes under Hatch-Waxman than it would otherwise.²¹⁸ The problem arises because there is now an additional benefit to settling: it reduces the number of generic rivals who will enter prior to patent expiration. This is a boon to both the patentee and the rivals who obtain settlements. Since the firms negotiate over profits afforded by this new benefit, bargaining possibilities expand. This enables the firms to agree on terms that restrain competition disproportionately, even if the same settlement design would not otherwise permit such a result. The rivals with credible challenge threats are willing to be restrained excessively — for example, to accept a disproportionately long delay period, or a disproportionately high royalty — in exchange for the opportunity to face a smaller competitive field upon entering the market.

This can be seen by comparing Tables 2 and 4. Recall that these tables correspond to pure delay settlements reached with 10 years remaining in the patent term, but differ in the number of prospective rivals; Table 2 contemplates a single rival, while Table 4 contemplates three. In both cases, the monopoly profit is 10 per year, while open competition gives each firm an annual profit of either 3 (in the single-rival case) or 1 (three rivals).²¹⁹ In the three-rival case, if the patent is 40% likely to be valid, we noted that litigation provides each rival with an expected profit of 6 over the remaining patent term. If all rivals obtain settlements, the only mutually-agreeable settlement outcome involved a four-year delay period, which is proportional. But now suppose that, due to Hatch-Waxman, only one rival has a credible challenge threat. Then, even though all three rivals would enter upon invalidation, the settlement ensures that only one of them will do so before the patent expires. Based on Table 2, this lone rival is willing to accept a delay period of up to 8 years, as this outcome provides the same profit that it expects to get through litigation.

As this illustrates, the problem arises simply because settlement results in a smaller competitive field than invalidation would have. It has nothing to do with the particular licensing restraint employed by the relevant settlements. As a consequence, the problem should not be attributed to the settlement arrangement itself. Instead, it should be

217. See Elhauge & Krueger, *supra* note 18, at 314; Hovenkamp & Lemus, *supra* note 89, at 33.

218. More accurately, this problem arises whenever the number of rivals who obtain settlements is smaller than the number who would have entered prior to expiration if the patent had been invalidated. This problem can arise from any phenomena that significantly undermine challenge incentives.

219. Total profits under open competition are lower in the three-rival case because there is greater competition when there are four firms in the market rather than two.

viewed as an institutional problem — namely that Hatch-Waxman undermines challenge incentives. This suggests that antitrust intervention is not appropriate so long as the settlement design is otherwise proportional.²²⁰ After all, there is no alternative arrangement that we could reliably expect to elicit better results in this regulatory environment.

B. Litigation Costs

When litigants settle, one element of the resulting gains from trade is simply the avoidance of litigation costs. The parties will bargain over the distribution of these cost savings, just as all contracting parties bargain over how to divide the gains from trade. In the patent settlement context, this bargain may in turn influence the bargain over the licensing restraint.

This is easily seen by considering what each party can agree to, which determines the possible outcomes of settlement negotiations. For instance, suppose the firms have decided to rely on a pure delay settlement and are bargaining over the delay period. If there are no litigation costs then, as we found earlier, they can only agree on the exactly proportional outcome. Let us suppose that this involves a four-year delay period. But with positive litigation costs, the expected profits from litigation fall for both firms. That means that the rival will agree to a slightly longer delay period — say, five years — whereas the patentee will agree to a slightly shorter delay period — say, three years. As such, the set of mutually-agreeable outcomes expands in both directions: relative to the case without litigation costs, the settlement outcome could be either more or less restrictive. It is still possible that they will agree on the proportional four-year delay period, but this is no longer the only possible bargaining outcome.²²¹

This ambiguous impact on the bargain over the licensing restraint is the first relevant effect of litigation costs. The second is that, if the patentee is sufficiently likely to win in court, then the rival's expected profit from litigation may become negative. In this case, the rival no longer has a credible litigation threat, since it expects to lose money if it litigates. As a consequence, litigation costs may undermine the rival's incentive to challenge the patent.²²²

220. See, e.g., *City of Pittsburgh v. W. Penn Power Co.*, 147 F.3d 256, 265 (3d Cir. 1998) (noting antitrust liability is inappropriate where alleged harm “did not flow from the defendants’ conduct, but, rather, from the realities of the regulated environment”).

221. If the firms had instead used a disproportionate licensing restraint, bargaining possibilities would similarly expand, but this time they would be centered around a disproportionate restraint magnitude rather than the proportional one. Thus, we would still expect private bargaining to result in a disproportionate result.

222. A corollary is that we cannot hope to encourage more patent challenges by permitting settlements that restrain competition disproportionately. Whether the rival has a credible challenge threat — which is what determines whether any challenge or settlement will occur —

Importantly, these two considerations do not depend on the settlement format at all. The first arises under any settlement design, and the second has nothing to do with settlement design, since it depends entirely on litigation expectations. The result is that litigation costs have very limited antitrust relevance. There is simply no way to avoid either possible effect of litigation costs on private bargaining. The best we can do is to limit the firms to proportional settlement formats, for at least in this case we know that bargaining possibilities will be centered around the proportional outcome, rather than a more restrictive one.

C. Cross-Licensing Settlements

The foregoing analysis focuses on settlements between a single patent holder and one or more rivals. But in practice, many settlements resolve disputes between two or more patent holders who compete in a product market,²²³ with each party's portfolio including some patents relevant to the dispute.²²⁴ In these cases, the settlement will typically effect cross-licensing, rather than a one-way transfer of technology rights.²²⁵ But the underlying antitrust tension is the same as in the single-patentee case. In order for the firms to mutually agree, it may be necessary for the settlement to restrain competition to some extent, but the firms always prefer to restrain competition to the cartel level so long as they can amicably divide the ensuing profits.

However, the tension is less pervasive in this context than in the single-patentee case, because it is less likely that the parties will be unable to mutually agree absent some restraint on competition.²²⁶ This is

depends only on litigation expectations. For instance, even if the rival *would* be allowed to obtain a reverse payment settlement, the patentee would not agree to offer such payment if it thinks the rival's litigation threat is just a bluff.

223. In some cases, the patentees are not product market competitors, but rather operate as licensors in a common technology space. In such cases, their preferred arrangement is typically to form a patent pool, which will centralize the licensing of both firms' patents. The most important question is whether the firms' patents are substitutes or complements from the perspective of licensees. Pooling is generally efficient in the latter case, but raises serious antitrust concerns in the former case. This point is well-understood in the literature. *See, e.g.,* CARL SHAPIRO, *INNOVATION POLICY AND THE ECONOMY* 119–50 (Adam Jaffe et al. eds., 2001); Richard Gilbert, *Antitrust for Patent Pools: A Century of Policy Evolution*, 2004 STAN. TECH. L. REV. 3 (2004); Josh Lerner & Jean Tirole, *Efficient Patent Pools*, 94 AM. ECON. REV. 691 (2004).

224. *See, e.g.,* *Standard Oil Co. (Ind.) v. United States*, 283 U.S. 163, 166 (1931); *United States v. Line Material Co.*, 333 U.S. 287, 289 (1948); *E. Bement & Sons. v. Nat'l Harrow Co.*, 186 U.S. 70, 80 (1902); *United States v. Singer Mfg. Co.*, 374 U.S. 174, 196–97 (1963).

225. This is so if the patents cover complementary or independent technologies; in either case, each firm may have an interest in using its rival's invention alongside its own. By contrast, if the relevant technologies were strictly substitutes, each firm would commercialize only one of them. In this case, the firms would have little incentive to challenge each other's patents as invalid, since that would just invite third-party competition. *See* Jay Pil Choi, *Live and Let Live: A Tale of Weak Patents*, 3 J. EUR. ECON. ASS'N 724, 726–29 (2005).

226. Hovenkamp & Lemus, *supra* note 7, at 29.

because the settlement occurs in the shadow of countervailing infringement actions; litigation could result in either or both firms being enjoined from using the other's patented technology. As a result of this comparative symmetry, the profits expected to accrue from litigation do not necessarily skew sharply in one party's favor as they frequently do in the single-patentee case.²²⁷ If the firms' bargaining positions are sufficiently symmetric, then they result in a wash, in which case the settlement need not restrain competition in order to be mutually-preferred to litigation.²²⁸

However, it may be that one party's litigation threat is much more powerful, leaving this "advantaged firm" with considerably more leverage. For instance, it may be that the advantaged firm's patent covers a technology that is near-essential to the relevant product category, whereas the other firm's technology is merely an optional enhancement. Under such circumstances, it may be that the firms cannot mutually agree without some restraint on competition — in particular, a restraint on the disadvantaged firm. Critically, however, it is never necessary for *both* firms to be restrained. Aside from being unnecessary, this leads to the same Pareto problem highlighted in the foregoing discussion of counter-restraints.²²⁹

Intuitively, the advantaged firm maintains a position much like the lone patentee in a settlement involving a single patent holder, and the disadvantaged firm's position is akin to that of a non-patentee rival. In fact, this underpins an important point: restrictive cross-licensing settlements can be evaluated using the same analytical approach applied to settlements involving a one-way transfer of licensing rights.²³⁰ The relevant aspects of settlement design are no different. There is some licensing restraint imposed on the disadvantaged firm; by default, its distributive properties will determine what the firms can agree on and how it compares to the expected result of litigation. But, as before, the settlement might also include a reverse payment (one paid by the unrestrained firm to the restrained one) or a counter-restraint (leaving both firms restrained). And, for the same reasons identified earlier, each of the latter two possibilities will always destroy proportionality. For these reasons, it is not necessary to develop a distinct economic rubric for evaluating cross-licensing settlements.

227. In the single-patentee case, litigation could leave the patent holder with a monopoly, but there is no chance of the rival obtaining one. Thus, expected profits from litigation skew in the patentee's favor to an extent that increases with the patentee's odds of prevailing.

228. For instance, the firms might just agree to exchange (unrestricted) royalty-free licenses.

229. See *supra* Section III.C.

230. Hovenkamp & Lemus, *supra* note 7, at 29–30 (demonstrating that the analysis of cross-licensing settlements collapses to the same proportionality analysis applied in the single-patentee case).

V. REFRAMING THE ANTITRUST QUESTION

The design-focused approach involves a reframing of the antitrust analysis. It does not attempt to discern the likelihood litigation would have resulted in an injunction, by proxy or otherwise. In this way, the design-focused approach allays the *Actavis* dissenters' primary objection to the majority's decision: that it would embed issues of pure patent law within the merits of the antitrust claim.²³¹ Instead, the relevant question is whether the firms' settlement is the kind of agreement that will induce reasonably proportional outcomes through private bargaining. In a word: will it keep the firms honest?

This approach ultimately focuses attention on how the settlement format affects the bargaining process. The single most important consideration is whether the settlement design maintains inter-party adversity as to the extent to which competition is restrained. As noted earlier, if private bargaining is to elicit proportional effects, it is critical that the patentee's rival is genuinely striving to be restrained as little as possible.²³² Maintaining adversity in the negotiation process should be viewed as a special embodiment of antitrust's more general objective of promoting competition. A horizontal patent settlement is a bargain over a restraint on inter-party competition. And, just as antitrust disfavors agreements that undermine inter-party competition directly, it ought to disfavor arrangements designed to slacken the settlement tug-of-war.

A. Administration

This section explains how the proportional-effects rule can be effectively administered under the rule of reason. This analysis is broken into two parts. The first addresses cases in which the challenged settlement involves a reverse payment or counter-restraint. I then turn to settlements that do not include either such provision, but which may nevertheless raise antitrust concerns based on the nature of the licensing restraint.²³³

231. *FTC v. Actavis, Inc.*, 570 U.S. 136, 163 (2013) (Roberts, J., dissenting) (“[W]e’re not quite sure if the patent is actually valid, or if the competitor is infringing it. But that is always the case, and is plainly a question of patent law. The majority, however, would assess those patent law issues according to ‘antitrust policies.’”).

232. See *supra* Section III.C.3.

233. Recall from Section III.B., that some licensing restraints always lead to disproportionate results, although the problem is not as pronounced as with reverse payments or counter-restraints.

A number of authors have proposed that reverse payments should be presumptively illegal.²³⁴ While the *Actavis* decision declined to apply such a rule,²³⁵ some commentators have predicted that the courts will nevertheless converge toward a de facto presumption of illegality.²³⁶ As discussed later, a number of recent cases give us serious reason to doubt that all lower courts will apply *Actavis* that way.²³⁷ Nevertheless, I agree with these authors that reverse payments are presumptively anticompetitive and should be adjudicated as such. Below I provide a more detailed proposal to that effect. Since counter-restraints undermine bargaining in the same way, they ought to receive the same treatment.

The *Actavis* majority held that pay-for-delay settlements are to be evaluated under the rule of reason.²³⁸ But such a decision does not preclude lower courts from maintaining a low bar for establishing a violation, which may amount to a de facto presumption of illegality.²³⁹ As the majority emphasized, “[t]here is always something of a sliding scale in appraising reasonableness” of conduct challenged under the antitrust laws; “the quality of proof required should vary with the circumstances.”²⁴⁰ The opinion further noted that lower courts may take advantage of this flexibility to ensure the rule of reason analysis is no more or less demanding than is reasonably necessary to appraise defendants’ conduct.²⁴¹ As such, the Court contemplated that an abbreviated rule of reason analysis is appropriate when the likelihood of anticompetitive effect is readily apparent.

234. See, e.g., Carrier, *supra* note 18, at 67; Thomas F. Cotter, *FTC v. Actavis, Inc.: When is the Rule of Reason Not the Rule of Reason?* 15 MINN. J.L. SCI. & TECH. 41, 42 (2014); cf. Carl Shapiro, *Antitrust Analysis of Patent Settlements Between Rivals*, 17 ANTITRUST 70, 72 (2003) (arguing that reverse payments are harmful). But see Daniel A. Crane, “*Per Se Illegality for Reverse Payment Settlements?*” Review of “*Unsettling Drug Patent Settlements: A Framework for Presumptive Illegality*,” 61 ALA. L. REV. 575, 575 (2010) (criticizing arguments for antitrust rules focusing on reverse payment settlements as a category).

235. *Actavis*, 570 U.S. at 158–59 (noting pay-for-delay settlements are subject to rule of reason review, not presumptive or per se illegality).

236. Cotter, *supra* note 234, at 43 (“[T]he Court appears to have all but in name adopted the presumptive illegality approach it purported to reject.”); Aaron Edlin et al., *Activating Actavis*, 28 ANTITRUST 16, 17 (2013) (“[T]he Court also made clear that a ‘long form’ rule of reason was not necessary, and in particular that both anticompetitive effect and market power could be inferred from large reverse payments themselves.”).

237. See discussion *infra* Section V.B.

238. 570 U.S. at 159.

239. See Cotter, *supra* note 234, at 43.

240. 570 U.S. at 159 (quoting *Cal. Dental Ass’n v. FTC*, 526 U.S. 756, 780 (1999)).

241. *Id.* at 159–60 (“As in other areas of law, trial courts can structure antitrust litigation so as to avoid, on the one hand, the use of antitrust theories too abbreviated to permit proper analysis, and, on the other, consideration of every possible fact or theory irrespective of the minimal light it may shed on the [merits]. We therefore leave to the lower courts the structuring of the present rule-of-reason antitrust litigation.”).

The rule of reason is a burden shifting framework for evaluating conduct that is not transparently anticompetitive.²⁴² First, the plaintiff must make a prima facie case: it must show that the defendant's conduct is likely to injure competition in a relevant market.²⁴³ If successful, the burden shifts to the defendant, who must show that there is a legitimate procompetitive purpose for its conduct.²⁴⁴ If the defendant establishes some such efficiency, the burden would then shift back to the plaintiff to show that there is a "less restrictive alternative" (LRA) to the defendant's conduct.²⁴⁵ An LRA is an alternative practice that: (a) would achieve the same procompetitive efficiency while having a milder effect on competition; and (b) would be reasonably commercially viable for the defendant.²⁴⁶ If the plaintiff fails to establish the existence of an LRA, then the dispute proceeds to the final stage: the balancing of pro- and anticompetitive effects.²⁴⁷

In the context of horizontal patent settlements, the plaintiff should have little difficulty making a prima facie case, since the settlement plainly restrains competition between the defendants.²⁴⁸ Similarly, the defendants should be able to meet their obligation at the second stage. As the *Actavis* majority noted, avoiding the lengthy and expensive litigation process is a legitimate procompetitive effect of settlement.²⁴⁹

Considerations of settlement design are most naturally addressed in the third stage's analysis of LRAs. Specifically, the relevant LRA would be an alternative settlement format — namely one that induces

242. See *United States v. Addyston Pipe & Steel Co.*, 85 F. 271, 282–83 (6th Cir. 1898), *aff'd*, 175 U.S. 211 (1899) (establishing the rule of reason); see also Roger D. Blair & Daniel Sokol, *The Rule of Reason and the Goals of Antitrust: An Economic Approach*, 78 ANTITRUST L.J. 471, 471 (2012) (discussing application of rule of reason to conduct that is not "presumptively unlawful"). Conduct that is transparently anticompetitive, such as naked price-fixing, is subject to a per se rule rather than the rule of reason.

243. See, e.g., Gregory J. Werden, *Cross-Market Balancing of Competitive Effects: What is the Law and What Should It Be?*, 43 J. CORP. L. 119, 132–40 (2017) (providing summary of the burden-shifting stages). The defendant's market power — typically a necessary condition for finding a violation — is also evaluated in the first stage.

244. In this stage, the defendant need not show that this procompetitive effect negates or outweighs the anticompetitive threat established in the first stage. It need only show that its conduct has *some* procompetitive effect. See, e.g., *Geneva Pharm. Tech. Corp. v. Barr Labs. Inc.*, 386 F.3d 485, 507 (2d Cir. 2004) ("If the plaintiffs satisfy their initial burden, the burden shifts to the defendants to offer evidence of the pro-competitive effects of their agreement.") (citing *Capital Imaging Assocs., P.C. v. Mohawk Valley Med. Assocs., Inc.*, 996 F.2d 537, 543 (2d Cir. 1993)).

245. For a rich discussion of LRAs, see generally C. Scott Hemphill, *Less Restrictive Alternatives in Antitrust Law*, 116 COLUM. L. REV. 927 (2016).

246. *Id.*

247. The plaintiff retains the burden of persuasion at the final stage, given its underlying obligation to prove a violation by a preponderance of evidence. See Werden, *supra* note 243, at 138–39.

248. See Cotter, *supra* note 234, at 45 (noting plaintiff will carry its initial burden because "there is obviously a contract that potentially restrains trade; that much is indisputable").

249. *FTC v. Actavis, Inc.*, 570 U.S. 136, 153 (2013) ("We recognize the value of settlements and the patent litigation problem.").

proportional effects through private bargaining. Consistent with this, the foregoing analysis supports a presumption that a settlement including a licensing restraint alone is an LRA to any settlement format that further includes a reverse payment or counter-restraint. Perhaps the most logical LRA would be a settlement arrangement that includes the same kind of licensing restraint as the challenged settlement, but without the other problematic provisions. For instance, in a pay-for-delay case, the presumptive LRA would be a pure delay settlement.

A proportional settlement gives the firms precisely the same profits they expect to garner through litigation, in addition to the savings from avoiding litigation. Thus, by definition, it provides a resolution that is viable but no more restrictive than necessary. This is precisely the logic behind LRAs generally. To further reinforce this point, recall that in almost all other litigation contexts, parties will naturally settle on terms that leave them in the same positions they expect to maintain post-litigation.²⁵⁰ This is relevant to establish that proportional settlements are sufficiently practical.²⁵¹ As Scott Hemphill notes: “The LRA ought to be practically rooted in commercial experience. . . . One useful indicia of practicality is that the alternative has been implemented by this or other firms in similar circumstances.”²⁵²

Under this standard, the evidence for the viability of proportional settlements is overwhelming. By compelling the parties to settle on terms that reflect their litigation expectations, a proportional settlement design merely imposes the same bargaining constraint that all other private litigants face by default. Given the ubiquity of private settlements generally, one cannot plausibly argue that this would be impracticable. This supports a presumption that an LRA exists in any cases involving a reverse payment or counter-restraint. The burden then falls on the defendant to rebut that presumption.

In fact, there is one settlement context — very far removed from patents or antitrust — that provides a nice illustration. It is very common for neighboring landowners to dispute the location of the boundary that separates their properties. Frequently the parties are uncertain as to the “true” location, and so they frequently elect to prescribe a certain boundary by mutual agreement.²⁵³ It is easy to see that with respect to

250. See *supra* Section II.B.1.

251. See, e.g., *M & H Tire Co. v. Hoosier Racing Tire Corp.*, 733 F.2d 973, 987 (1st Cir. 1984) (rejecting proposed LRA for being “more hypothetical than practical”); U.S. DEP’T OF JUSTICE & FTC, HORIZONTAL MERGER GUIDELINES § 10 (2010) (noting proposed alternatives should be “practical” and not “merely theoretical”).

252. Hemphill, *supra* note 245, at 985.

253. *Young v. Blakeman*, 153 Cal. 477, 481 (1908) (“[T]he rule has been established that when such owners, being uncertain of the true position of the boundary so described, agree upon its true location . . . such line becomes, in law, the true line . . . regardless of the accuracy of the agreed location . . .”); MITCHELL J. WALDMAN, FLORIDA JURISPRUDENCE § 45 (2d ed. 2018) (“Where the boundary between contiguous lands is uncertain or disputed, the owners of such lands may agree upon a certain line as the permanent boundary line.”).

the underlying bargaining problem, these settlements are highly analogous to pure delay settlements. The parties choose the boundary in the shadow of litigation, so neither one will agree to give up significantly more ground than it expects it would obtain through litigation. Thus, their ultimate agreement will reflect their expectations about the boundary that would be prescribed by a court.

A final point is that, in order to rebut an LRA presumption, the defendant firms cannot simply rely on bald assertions that they would not have settled under the LRA format because they didn't want to, or because one of them asserted its refusal to do so. Such was the bewildering approach taken in *In re Wellbutrin XL Antitrust Litigation Indirect Purchaser Class*,²⁵⁴ where the district court held that a pure delay settlement was not a viable alternative because the patentee's rival "unwaveringly refused to settle . . . unless the settlement contained a no authorized generic [provision]."²⁵⁵ This suggests that two firms can avoid liability for an anticompetitive agreement by asserting that they were determined to enter into one.²⁵⁶ That leads to the perverse implication that the relevant antitrust limits should be fashioned around the preferences of the firms, which is exactly backwards. Such limits are necessary in the first place only because the firms have a strong preference for stifling competition.

It should be clear that the relevant question is not whether the firms *want* to employ the LRA. Of course they don't — it is less restrictive and therefore less profitable. Rather, the relevant question is whether the LRA settlement format is the kind of agreement that can permit two competing parties to resolve a patent dispute and thereby provide the relevant efficiency, avoiding litigation. This assessment should be driven by objective economic analysis, not bald assertions by the firms. To rebut a presumptive LRA, the firms must show that it is not capable of providing a profit allocation that emulates the expected result of litigation. Thus, for example, a non-restrictive settlement format is not a viable LRA; it preserves the competitive level of total profits, whereas litigation affords super-competitive profits in expectation.²⁵⁷ But a proportional settlement format like pure delay does not suffer this deficiency. Just like neighboring landowners settling on a boundary line, pure delay provides the flexibility to emulate the parties' litigation expectations, whatever they happen to be.

254. 133 F. Supp. 3d 734 (2015).

255. *Id.* at 757. Recall that a no authorized generic provision is a counter-restraint.

256. Soter, *supra* note 23, at 1336 ("Such a rule tells defendants that all they need to do to avoid liability is to insist in settlement talks that the only agreement they would make is an [anticompetitive] one.").

257. See discussion *supra* Section II.B.

Now consider the case where the firms' settlement includes only a licensing restraint on the patentee's rival. As explained earlier, all proportional settlement agreements — such as a pure delay or an ordinary royalty obligation — fit this description.²⁵⁸ But not all licensing restraints are proportional. Some problematic restraints, such as restrictions on price or output, compel a disproportionate result, because they make it impossible for the firms to mutually agree on anything less.²⁵⁹

When the agreement includes only a licensing restraint, the firms cannot bargain separately over the size and distribution of total profits. Instead, these two variables are jointly determined by the magnitude with which the restraint is applied (e.g. the length of a delay period). As the magnitude increases, total profits rise and the rival's profit share falls, so that the distribution skews increasingly in the patentee's favor.²⁶⁰ Absent a reverse payment or counter-restraint, the firms cannot modify these profit distributions, which is why they generally cannot agree to restrain competition all the way to monopoly (which would leave the rival with zero profits).²⁶¹ Instead, there is a genuine tug-of-war over the intensity with which competition is restrained, since this necessarily governs each firm's payout.

Evaluating the differences among licensing restraints is where the economic analysis becomes most complicated,²⁶² and the technical details are outside the scope of this paper. But the important takeaway is that the choice of licensing restraint determines proportionality by virtue of how it distributes profits. Here too the analysis of LRAs is the most natural place to park these considerations. If the defendant firms employ a potentially problematic restraint — such as a floor on the rival's price²⁶³ — a plaintiff may reasonably posit an alternative restraint as an LRA. For example, as noted earlier, royalties generally elicit proportional effects, as do restraints on entry. Since such restraints are common, there is reason to believe that they are generally viable. If the defendants cannot articulate why this is not true within the context of their dealings,²⁶⁴ the plaintiff ought to prevail on an LRA theory.

258. *See id.*

259. *See supra* notes 157–172 and accompanying text.

260. *See supra* Tables 1–3.

261. The rival would not have a credible threat to litigate unless it expected it could obtain a positive profit through litigation, in which case it will not settle for nothing.

262. Hovenkamp & Lemus, *supra* note 7, at 20–21 (introducing a tool called a “substitution rate” that can be used to appraise a licensing restraint on proportionality grounds).

263. *See supra* notes 150–163 and accompanying text (discussing price restraints).

264. A restraint that is proportional in most contexts could become nonviable under certain market conditions. For instance, in a market with highly aggressive price competition, a royalty may be nonviable since the rival may be unable to compete at all if it is at a material cost disadvantage. This may explain why pharmaceutical patent settlements tend to rely on entry restraints rather than on royalties.

B. Demystifying Causation

This alternative framing of the antitrust question would help to resolve the post-*Actavis* confusion that has pervaded among some courts. Since the Supreme Court issued the decision in 2013, there have been a number of judgments entered in pay-for-delay cases. The results have been mixed.²⁶⁵ There is a divergence among courts as to: (a) what showings are required to establish a prima facie violation; and (b) what a private plaintiff must show in order to satisfy the causation element of antitrust standing.²⁶⁶ To establish standing, some courts have required private plaintiffs to assert a particular theory as to how generic entry would have occurred earlier — *definitively* earlier, not earlier in expectation — but for the settlement.²⁶⁷ This simultaneously (1) conflates the causation element of private standing with the merits and (2) overstates the plaintiff's burden of production with respect to the alleged violation.

In practical effect, these courts require private plaintiffs to show either that the patent holder would have lost in hypothetical patent litigation, or else that the parties would have entered into a less restrictive settlement without a reverse payment (i.e. a pure delay settlement).²⁶⁸ As to the former option, in *In re Nexium (Esomeprazole) Antitrust Litigation*,²⁶⁹ the court suggested that plaintiffs were required to demonstrate invalidity or noninfringement of the relevant patent in order to prevail.²⁷⁰ In *Wellbutrin*, the Third Circuit largely rejected the Supreme Court's contention that a reverse payment can typically support a finding that the patent is likely invalid.²⁷¹ In fact, this was just one of several

265. For a richer discussion of this post-*Actavis* divergence, see Soter, *supra* note 23, at 1313–28.

266. See *id.* at 1311 (noting that in some recent private actions, courts have conflated causation with substantive merits of antitrust claim). This concern would not arise in a public enforcement action, since the antitrust agencies have statutory standing. See Hovenkamp, *supra* note 81, at 54 (discussing causation in private antitrust actions).

267. These cases suggest that plaintiffs must choose from three possible theories: (a) the generic would have entered “at risk” (i.e. without securing a license); (b) the generic would have prevailed in patent litigation; or (c) the generic would have obtained an earlier entry date via a payment-free (“pure delay”) settlement. See, e.g., *In re Opana ER Antitrust Litig.*, 162 F. Supp. 3d 704, 720 (N.D. Ill. 2016); Soter, *supra* note 23, at 1314.

268. See 162 F. Supp. 3d at 720. See also Soter, *supra* note 23, at 1315.

269. 842 F.3d 24 (1st Cir. 2016).

270. *Id.* at 62 (“Plaintiffs respond that they should not have to prove patent invalidity or noninfringement . . . but to no avail.”).

271. *In re Wellbutrin XL Antitrust Litig. Indirect Purchaser Class*, 868 F.3d 132, 168 (3d Cir. 2017). Compare *id.* (“While the size of a reverse payment may have some relevance . . . it is far from dispositive.”), with *FTC v. Actavis*, 570 U.S. 136, 157 (“It is normally not necessary to litigate patent validity An unexplained large reverse payment itself would normally suggest that the patentee has serious doubts about the patents survival.”).

points at which the *Wellbutrin* court rebuffed express pronouncements made by the *Actavis* majority.²⁷²

Fortunately, some other courts have correctly interpreted and applied the *Actavis* decision, espousing the proportional-effects rule as the relevant standard for enforcement.²⁷³ But the lingering confusion among some courts is nevertheless concerning. It seems to center on an underlying discomfort with the “Actavis Inference” — the reliance on a large reverse payment as a signal that the relevant patent is likely invalid or noninfringed. Some courts do not find this to be a reliable proxy,²⁷⁴ and instead regard a case-within-a-case analysis of the relevant patent as the only sufficient means of establishing a violation.²⁷⁵

The design-focused approach would go a long way in resolving this post-*Actavis* confusion. It eliminates the need to rely on the Actavis Inference. Hence, courts need not fuss over what can be reasonably inferred from a particular reverse payment in a particular case. Instead, the plaintiff can simply show that the settlement format undermines the bargaining process, naturally inducing the firms to agree on terms that restrain competition disproportionately. *That* is the violation.²⁷⁶ This is consistent with how other kinds of antitrust issues are adjudicated. In general, for a private plaintiff to prevail on the merits, it must show that it was (or likely will be) harmed — to one extent or another — by conduct violating the antitrust laws.²⁷⁷ This is sufficient to obtain an injunction. It is only the availability of damages that requires a private plaintiff’s injury to be reasonably quantified.²⁷⁸

272. See Michael Carrier, *The Curious Case of Wellbutrin: How the Third Circuit Mistook Itself for the Supreme Court*, 103 CORNELL L. REV. 74, 75 (2017) (criticizing *Wellbutrin* court for eschewing numerous aspects of *Actavis* decision).

273. *In re Cipro Cases I & II*, 348 P.3d 845, 863–64 (Cal. 2015) (“*Actavis* embraces the insights of Professor Carl Shapiro and others that the relevant benchmark in evaluating reverse payment patent settlements should be no different from the benchmark in evaluating any other challenged agreement: What would the state of competition have been without the agreement?”).

274. For instance, the *Wellbutrin* court was persuaded by some amici that the reverse payment would not signal a likelihood of invalidity if the patent holder is risk-averse. 868 F.3d at 168. *But see infra* Section V.C. (criticizing risk aversion argument).

275. See, e.g., *Soter*, *supra* note 23, at 1317–18 (noting that some courts seem to require showing of invalidity or noninfringement by a preponderance of evidence).

276. See *FTC v. Actavis, Inc.*, 570 U.S. 136, 157 (2013) (“[T]he payment (if otherwise unexplained) likely seeks to prevent the risk of competition. And, as we have said, *that consequence constitutes the relevant anticompetitive harm.*”) (emphasis added).

277. 15 U.S.C. § 26 (2012) (stating private plaintiff “shall be entitled to sue for and have injunctive relief . . . against threatened loss or damage by a violation of the antitrust laws”). The alleged harm must also constitute an “antitrust injury,” although this point is ancillary to the present discussion. See *Cargill, Inc. v. Monfort of Colorado, Inc.*, 479 U.S. 104, 122 (1986).

278. See *Blue Cross and Blue Shield United of Wisconsin v. Marshfield Clinic*, 152 F.3d 588, 591 (7th Cir. 1998) (“[Where private plaintiff] has failed to come up with evidence that would authorize an award of damages for the [antitrust violation], this does not justify with-

C. Evaluating Potential Objections

This section considers a number of potential objections to this paper's proposals, with emphasis on those invoked most frequently in the literature. Most of them come from commentaries on reverse payment, which has dominated the settlement literature in recent years. But the same arguments could be applied to other settlement provisions garnering antitrust scrutiny.²⁷⁹

1. What if the Patentee is Risk-Averse?

A number of commentators have proposed that reverse payments may not be problematic because they may merely reflect that the patentee is risk-averse.²⁸⁰ In *Wellbutrin*, the Third Circuit found this suggestion sufficiently persuasive to rebut the Actavis Inference.²⁸¹ An agent is said to be risk-averse when the value it imputes to a probabilistic payout is strictly less than the expected value of such payout.²⁸² In effect, risk-averse agents impute a psychological “cost” to uncertainty itself. With respect to the patent settlement debate, the argument is that a risk-averse patentee would be willing to make a payment just to avoid the uncertainty of litigation, and hence such payment may not result in competition being restrained disproportionately.²⁸³

In industrial organization, the branch of economics applied in antitrust, firms are almost always treated as risk neutral, meaning that they maximize the expected value of profits. Such firms will not sacrifice expected profits for the sake of avoiding uncertainty. Some commentators have offered cogent skepticism of the suggestion that corporate patentees are likely to be risk-averse.²⁸⁴ However, even if we concede that

holding an injunction — rather the contrary. Inadequacy of [damages] is normally . . . a prerequisite to the entry of an injunction. . . . And a common reason why the damages remedy is inadequate is that the plaintiff is unable to quantify the harm”)

279. For instance, a no-authorized-generic commitment (which is a counter-restraint) has received significant attention in the aftermath of *Actavis*. See, e.g., Edlin et al., *supra* note 18, at 585.

280. See, e.g., Crane, *supra* note 1, at 704; Schildkraut, *supra* note 194, at 1058.

281. See *In re Wellbutrin XL Antitrust Litig. Indirect Purchaser Class*, 868 F.3d 132, 168–69 (3d Cir. 2017).

282. For example, if a lottery ticket has a 5% chance of being worth \$100, then a risk neutral party values the ticket at $5\% \times \$100 = \5 , whereas a risk averse party values it at some lesser amount — say, \$4.

283. See, e.g., Henry N. Butler & Jeffrey Paul Jarosch, *Policy Reversal on Reverse Payments: Why Courts Should Not Follow the New DOJ Position on Reverse Payment Settlements*, 96 IOWA L. REV. 57, 95 (2010) (noting risk-averse patentee “might be willing to pay a reverse payment simply to avoid the risks of losing everything at trial”); Schildkraut, *supra* note 194, at 1058; Robert D. Willig & John P. Bigelow, *Antitrust Policy Toward Agreements That Settle Patent Litigation*, 49 ANTITRUST BULL. 655, 666 (2004).

284. See, e.g., Elhauge & Krueger, *supra* note 18, at 311–12.

a patentee may be risk-averse, there are clear reasons to reject this as a basis for permitting reverse payments.

First, if a defendant is indeed risk-averse, it properly bears the burden of proving this. An antitrust plaintiff should not be forced to rebut allegations relating to the defendant's subjective attitude toward risk. As the Supreme Court has noted, the "ordinary rule . . . does not place the burden upon a litigant of establishing facts peculiarly within the knowledge of his adversary."²⁸⁵ This is in direct contrast to the approach taken in *Wellbutrin*, where the court regarded mere allusions to hypothetical risk aversion as sufficient to rebut concerns relating to a reverse payment.²⁸⁶

Second, even if the patentee is truly risk-averse, it does not follow that consumers ought to bear the costs of avoiding the psychological costs it attributes to uncertain litigation. The firms can still settle using a proportional settlement format. For example, under pure delay, a risk-averse patentee is willing to accept a delay period that is shorter than the proportional one.²⁸⁷ In fact, by expanding the bargaining possibilities, this actually makes settlement easier. As an accounting matter, this provides the patentee with lower profits than the expected result of litigation. But it is better than litigation from the patentee's perspective, since it attaches subjective value to the avoidance of uncertainty. Thus the settlement is still mutually agreeable. In principle, the patentee could issue a reverse payment in order to raise the delay period to the proportional level. But this effectively forces consumers to finance the patentee's insurance bill.

However, by far the most important flaw in the risk-aversion argument relates to the "Pareto problem" discussed earlier.²⁸⁸ Even if a reverse payment provides a means of offsetting hypothetical risk aversion, this function cannot be disentangled from the payment's profit-sharing function. And, whatever their attitudes toward risk, we can be sure that the firms are not *profit*-averse. To that end, they can always share in a larger profit by further increasing the reverse payment, and we should fully expect that they will do so. In fact, this is just one embodiment of a more general problem with arguments supporting reverse payments, which is discussed further below.

285. *United States v. New York, New Haven & Hartford R.R. Co.*, 355 U.S. 253, 256 n.5 (1957).

286. 868 F.3d at 168.

287. See Butler & Jarosch, *supra* note 283, at 95 (noting risk-averse patentee "might be willing to accept an entry date significantly before the [expectation benchmark]").

288. See *supra* Section III.C.3.

2. Appeals to Non-Pareto Hypotheticals

Aside from risk aversion, advocates for reverse payment legality cite a few other bargaining problems that could in principle be resolved with a payment, provided it is not made too large. Another such problem involves the possibility that the firms do not maintain the same expectations about litigation.²⁸⁹ Another hypothesized problem is that the patentee's rival may be on the brink of foreclosure, and thus unable to survive the proportional delay period without a reverse payment to keep it going.²⁹⁰

In the context of pay-for-delay agreements, Edlin et al. emphasize the importance of focusing not merely on what settlement outcomes are mutually preferred to litigation, but on the subset of such possibilities that the profit-maximizing firms might actually pick.²⁹¹ Focusing on Pareto efficient bargaining outcomes is critical, for these are the ones rational firms will agree on. To that end, as already explained, both reverse payments and counter-restraints lead *all* Pareto efficient bargaining outcomes to restrain competition to the profit-maximizing level, regardless of what odds the firms impute to litigation. This problem is categorical; even in conjectural scenarios where a limited payment is alleged to be efficient, the problem is still there.

This underscores the critical flaw in most arguments favoring the legality of reverse payments, namely that they appeal to hypotheticals that are economically innumerate. More specifically, they appeal to bargaining outcomes that are Pareto inferior.²⁹² Such arguments typically proceed as follows. First, the arguer asserts some hypothetical problem or friction that interferes with private bargaining in some way. Second, the arguer shows that there is some reverse payment — which has some maximal dollar value X — that would allow the firms to overcome this problem, while still ensuring the settlement is proportional.²⁹³ Third, there is an unstated but critical assumption, which is that the firms might actually limit the payment to X , notwithstanding that they

289. See, e.g., Butler & Jarosch, *supra* note 283, at 98–99 (arguing that reverse payments may be necessary to resolve a disagreement as to how litigation is likely to play out). Note that, as with risk-aversion, an antitrust plaintiff will typically have no capacity to disprove an assertion that the firms maintain asymmetric priors.

290. *Id.* at 98. This particular hypothesis is dubious on its own terms. If the rival is too poor to survive until the proportional entry date, then, in light of the high cost of patent litigation, it is unlikely that it would have a credible litigation threat, in which case there would not be any settlement in the first place.

291. Edlin et al., *supra* note 18, at 614 (“The distinction between feasible outcomes and actual equilibrium outcomes is fundamental to neoclassical economic analysis . . .”). In the context of settlement contracting, “equilibrium” refers to a settlement outcome that is Pareto efficient and mutually preferred to litigation.

292. *Id.*

293. See, e.g., Willig & Bigelow, *supra* note 283, at 688–89; Butler & Jarosch, *supra* note 283, at 96, 98–99; Kobayashi et al., *Actavis and Multiple ANDA Entrants: Beyond the Temporary Duopoly*, 29 ANTITRUST 89, 92 (2015).

could both share in a larger profit by using a larger payment.

Such arguments thus ask us to entrust self-interested firms to choose lesser profits over larger ones. This is like hoping that my drink won't spill if I turn it upside-down. It defies the natural forces of markets and private contracting. The problem is exacerbated by the fact that, aside from being extremely hard to verify in the first place, the hypothesized bargaining problems would be virtually impossible for a court or antitrust plaintiff to quantify. Realistically, a court has no capacity to say whether or not a given payment is larger than necessary to accommodate, say, a patentee's risk-aversion. All firms, including those who do not suffer the hypothesized bargaining problem, know this in advance. All of them will thus use the largest payment they expect they can get away with, while assuring the court that such payment is necessary to cure a serious contracting friction. If the courts permit this, the firms would be irrational not to exploit it for profit. As such, under any antitrust regime that permits reverse payments as a cure for hypothetical bargaining frictions, we should expect that essentially all such payments to facilitate disproportionate effects on competition.

It is true that the firms may be unable to settle in some cases, but only for the usual reasons that all private litigants may fail to settle. The parties to any dispute occasionally have divergent expectations or asymmetric information as to the plaintiff's likelihood of prevailing.²⁹⁴ Such possibilities are widely regarded as the most common explanation for failures to settle.²⁹⁵ A proportional settlement design puts the firms in the same positions as ordinary litigants and is thus vulnerable to the same possibility of a failure to settle. But, again, there is nothing special about this; no other area of litigation is immune to it either.

Further, as with risk aversion, an antitrust plaintiff probably has no realistic ability to disprove the defendants' assertion that they disagree about the odds of patent litigation, even if it is flatly untrue. Thus, as before, if the courts were to permit reverse payments as a solution to a divergent expectations problem, we should expect all firms to start relying on this as a pretextual basis for excessive restraints on competition. The result would be to undermine all horizontal settlements just to help avoid settlement failures in a tiny percentage of disputes.

Another line of argument suggests that alternative settlement formats, which do not include reverse payments and are generally viewed as benign, may be just as restrictive as pay-for-delay. Here the relevance of Pareto efficiency is more subtle. For instance, in a thoughtful

294. Asymmetric information means that only one party knows the correct probability that the plaintiff will win. See, e.g., Joel Waldfogel, *Reconciling Asymmetric Information Theories and Divergent Expectations Theories of Litigation*, 41 J.L. ECON. 451, 451 (1998).

295. See, e.g., *id.*; Kathryn Spier, *Litigation*, in THE HANDBOOK OF LAW AND ECONOMICS 259, 321 (Kenneth J. Arrow & Michael G. Intriligator eds., 2007) (noting that most literature on failure to settle focuses on either divergent expectations or asymmetric information).

2014 article, Dan Crane gives an example in which the rival is made to pay a per-unit royalty equal to the monopoly markup.²⁹⁶ Professor Crane rightly notes that this essentially preserves monopoly, assuming the rival is not a more efficient producer, since the rival will be unable to set a price at which it will make sales.²⁹⁷ But, in contrast to a reverse payment settlement, here the rival makes no money when it is fully excluded. As such, the royalty settlement would be agreeable to the rival only if the patent were virtually certain to be valid and infringed.²⁹⁸ But in that case the settlement is proportional and therefore not an antitrust concern. Indeed, as noted earlier, royalties generally elicit proportional effects.²⁹⁹

This is a useful reminder that proportional settlement formats do not preclude agreements that restrain competition substantially. Rather, they ensure that Pareto efficient bargaining possibilities are always commensurate with the expected result of litigation. By contrast, problematic settlement formats eliminate any link between Pareto efficiency and the firms' litigation expectations, because the former will always involve shared monopoly.

VI. CONCLUSION

Patent settlements between rivals have long been a source of confusion and debate in antitrust. Recently, courts and scholars have converged toward a concrete standard — the proportional-effects rule — under which liability turns on how the settlement's competitive effects compare to the expected result of litigation. It has been presumed that this requires some assessment of the likelihood that the relevant patent is valid and infringed. But in practice this requires either an onerous case-within-a-case analysis, or else relies on an evidentiary proxy that is absent in most kinds of settlements that have garnered antitrust scrutiny over the years. Complicating matters, even when such proxy is available, some courts have demonstrated an unwillingness to accord any weight to it, notwithstanding that this contradicts an express pronouncement by the Supreme Court.

This paper explains that antitrust law would benefit substantially by reframing the settlement inquiry so as to strip away issues of pure

296. Daniel Crane, *Actavis, the Reverse Payment Fallacy, and the Continuing Need for Regulatory Solutions*, 15 MINN. J.L. SCI. & TECH. 51, 57 (2014).

297. The patentee can set price slightly below the monopoly level, and the rival would not be able to match this without incurring a loss. This way the patentee gets the entirety of a near-monopoly profit rather than half of the full-monopoly profit.

298. Here we are overlooking the fact that if the rival does not expect to earn any profit through litigation, then it lacks a credible litigation threat and will not obtain a settlement in the first place.

299. See *supra* notes 148–152 and accompanying text.

patent law from the underlying violation. This builds upon recent economic scholarship in which Jorge Lemus and I demonstrate that a settlement's design is always what determines the proportionality (or lack thereof) of its ultimate competitive effects, no matter the particular odds that the patent is valid and infringed. In the present article, I apply these economic arguments to antitrust law and practice. Among other things, I emphasize that: (a) the design-focused approach can be administered much more practicably, reliably, and broadly than the prevailing patent-focused approach; (b) it is consistent with the Supreme Court's *Actavis* decision; (c) it simplifies the antitrust analysis by disentangling the relevant antitrust violation from the extent of the resulting harm; and (d) it clarifies a number of critical errors in arguments advocating against antitrust intervention in patent settlements. These points support a clear prescription for antitrust reform: evaluate the agreement, not the patent.