



Another Take on the Relevant Welfare Standard for Antitrust

What is the goal of antitrust law? **Jon Jacobson** analyzes the strengths and weaknesses of the various welfare-based standards that have been proposed to answer that question. After reviewing the alternatives, he concludes that antitrust's objective is to protect the competitive process, with anticompetitive effects best analyzed through the impact on market output.

Innovators, Implementers, and Two-sided Hold-up

Luke Froeb and **Mikhael Shor** contend that certain policies recently adopted by standard-setting organizations and confirmed by the U.S. Department of Justice are excessive means of addressing patent hold-up problems and may ultimately harm the incentives for innovation. They point in particular to policies basing RAND terms on the smallest salable patent practicing unit and curtailing the use of injunctive relief by holders of SEPs.

Reflections on the Debates Surrounding Standard-Essential Patents

Greg Leonard reviews the current economic thinking concerning certain aspects of standard essential patents. He considers whether hold-up is a real problem, to what extent terms reached during patent licensing negotiations are indicative of RAND terms, and whether a reasonable royalty for an SEP under a RAND commitment should be based on an ex-ante valuation of the SEP or ex-post valuation. The author also describes an increasing gap between U.S. case law and the current economic analysis in this area.

Paper Trail: Working Papers and Recent Scholarship

Editor **Bill Page** reviews a paper by **Jeremy Bertomeu**, **John Harry Evans III**, **Mei Feng**, and **Ayung Tseng** addressing the anticompetitive effect of information sharing among firms, using the automobile industry as a test of the generated hypotheses. Editor **John Woodbury** considers a paper by **Greg Werden** challenging the outcomes of retrospective studies of mergers and another paper by **Justin Johnson** offering a benign explanation for “loss leaders” when consumers are characterized by bounded rationality.

Another Take on the Relevant Welfare Standard for Antitrust

Jonathan M. Jacobson

As Robert Bork explained in *The Antitrust Paradox*:¹

Antitrust policy cannot be made rational until we are able to give a firm answer to one question: What is the point of the law—what are its goals. Everything else follows from the answer we give . . . Only when the issue of goals has been settled is it possible to frame a coherent body of substantive antitrust rules.²

There is widespread agreement today that Judge Bork's assessment was correct. Antitrust policy cannot be coherent unless its goals are understood. The problem, however, is that there remains no consensus on what those goals should be.

In this article, I identify a variety of standards that have been proposed since *The Antitrust Paradox* was published. I analyze some of their various strengths and weaknesses and propose a standard for further study and analysis: that the goal of antitrust is to protect the competitive process, with anticompetitive effects best analyzed through the impact on market output.

Welfare Standards

The Supreme Court has never articulated a specific welfare standard. But several different ones have been proposed. In most antitrust cases, the choice of welfare standard really does not matter, as the same results will hold regardless of the standard applied. The instances in which the selection matters, however, can be quite important, and the standard chosen invariably says much about the decision maker's basic philosophy of antitrust. The most significant include the following:

Total Welfare. "Total" welfare looks to measure the effect of a practice or transaction on the economic welfare of all participants in a market, including both producers and consumers. Put differently, it "refers to the aggregate value that an economy produces, without regard for ways that gains or losses are distributed."³ Among the many proponents of the total welfare standard are Professors Roger Blair and Daniel Sokol⁴ and senior government economist Kenneth Heyer.⁵

¹ ROBERT H. BORK, *THE ANTITRUST PARADOX* (1978).

² *Id.* at 50.

³ 1 PHILLIP AREEDA & HERBERT HOVENKAMP, *ANTITRUST LAW* ¶ 114a (4th ed. 2014).

⁴ Roger D. Blair & D. Daniel Sokol, *Welfare Standards in U.S. and E.U. Antitrust Enforcement*, 81 *FORDHAM L. REV.* 2497 (2013); *The Rule of Reason and the Goals of Antitrust: An Economic Approach*, 78 *ANTITRUST L.J.* 471 (2012).

⁵ Kenneth Heyer, *Welfare Standards & Merger Analysis: Why Not the Best?* (EAG Discussion Paper 06-8 U.S. Dep't of Justice Mar. 2006); see also Alan J. Meese, *Debunking the Purchaser Welfare Account of Section 2 of the Sherman Act: How Harvard Brought Us a Total Welfare Standard and Why We Should Keep It*, 85 *N.Y.U. L. REV.* 659, 690-98 (2010); Charles F. Rule & David L. Meyer, *An Antitrust Enforcement Policy to Maximize the Economic Wealth of All Consumers*, 33 *ANTITRUST BULL.* 677 (1988).

Most observers also have understood Judge Bork's references to "consumer welfare" to refer to a total welfare standard,⁶ but his approach is better understood somewhat differently as a standard under which the goal is allocative efficiency to the extent it does not interfere significantly with productive efficiency.⁷

Consumer Welfare. The consumer welfare standard equates with consumers' surplus in economic terms—technically, the difference between what consumers actually pay and what they would be willing to pay. To illustrate the principle, consider a merger of rival firms that both reduces their costs and gives them market power.⁸ If costs are reduced but prices to consumers still rise, the merger is viewed as benign under a total welfare standard if the cost reduction is greater than the price increase. But the same merger will fail the consumer welfare standard unless the cost decrease is such that prices to consumers remain the same or fall. The gains to the merging producers do not count; only the effect on consumer prices is relevant. This consumer welfare standard is the standard understood to be employed in practice by the federal enforcement agencies,⁹ and is supported by many observers including, most preeminently, Professor Steven Salop.¹⁰

This consumer welfare standard is the standard understood to be employed in practice by the federal enforcement agencies . . .

Consumer Choice. The relatively new "consumer choice" standard is based on the idea that the "range of options [available to consumers should not] be significantly impaired or distorted by anticompetitive practices."¹¹ The standard is not based on any specified number of options, and does not forbid all reductions in choice, but focuses instead on "conduct that artificially limits the natural range of choices in the marketplace."¹²

Multiple Goals. For much of the first century of U.S. antitrust enforcement, the courts made clear that at least one purpose of the antitrust laws was the protection of small business—the "small dealers and worthy men" praised in *Trans-Missouri*¹³ and the "small, locally owned businesses" of *Brown Shoe*.¹⁴ Following these precedents, many observers concluded that antitrust's goals included preserving a deconcentrated industry structure, dispersing economic power, pro-

⁶ E.g., Einer Elhauge, *Tying, Bundled Discounts, and the Death of the Single Monopoly Profit Theory*, 123 HARV. L. REV. 397, 437–38 (2009).

⁷ BORK, *supra* note 1, ch. 5. Gregory Werden describes Bork's view as "general equilibrium social welfare": "General equilibrium social welfare relates only to actual consumers; it is the welfare of the people who make up the society. In contrast, partial equilibrium consumer surplus does not directly relate to consumers in most antitrust cases because businesses most often are the sellers and the buyers in the relevant market." Gregory J. Werden, *Antitrust's Rule of Reason: Only Competition Matters*, 79 ANTITRUST L.J. 713, 723 (2014). Werden provides a cogent explanation of Bork's use of the phrase "consumer welfare," and why it was not misleading. *Id.* at 718–23.

⁸ For the classic exposition of this example, see Oliver E. Williamson, *Economics as an Antitrust Defense: The Welfare Tradeoffs*, 58 AM. ECON. REV. 18 (1968).

⁹ See AREEDA & HOVENKAMP, *supra* note 3, ¶ 114b; U.S. Dep't of Justice & Fed. Trade Comm'n, Horizontal Merger Guidelines § 10 (2010).

¹⁰ E.g., Steven A. Salop, *Question: What Is the Real and Proper Antitrust Welfare Standard? Answer: The True Consumer Welfare Standard*, 22 LOYOLA CONSUMER L. REV. 336 (2010).

¹¹ Robert H. Lande, *Consumer Choice as the Ultimate Goal of Antitrust*, 62 U. PITT. L. REV. 503, 503 (2001).

¹² *Id.* at 503–04.

¹³ *United States v. Trans-Missouri Freight Ass'n*, 166 U.S. 290, 323 (1897).

¹⁴ *Brown Shoe Co. v. United States*, 370 U.S. 294, 344 (1962); *accord*, *United States v. Aluminum Co. of Am. (Alcoa)*, 148 F.2d 416, 427 (2d Cir. 1945).

moting fairness in economic dealings, and providing competitive market structures to reduce the need for governmental control.¹⁵

Competitive Process. A fifth option is the competitive process standard articulated by Gregory Werden and others.¹⁶ Under this approach, practices and transactions that interfere with competition as a process would be prohibited, focusing only on economic effect, but *without* focusing on any particular welfare standard. Practices that do not impair the competitive process would not be prohibited, even if there is some negative impact on consumer surplus.

Assessing the Alternatives

In the years following the passage of the Sherman Act, the protection of small business and related non-economic goals were at the forefront of antitrust enforcement. This was consistent with the original intent of the Sherman and Clayton Acts, later efforts to rewrite that history notwithstanding.¹⁷ But starting with *Sylvania*¹⁸ and *Brunswick*¹⁹ in 1977, only economic goals have mattered, and no one expects that to change.

The question today is what the standard should be in assessing the *economic* consequences of a practice or transaction. As mentioned, the Supreme Court has never articulated an answer. And Gregory Werden has demonstrated ably that efforts to tease a particular welfare standard out of the Supreme Court's opinions invariably fail.²⁰ The Court's references to a "consumer welfare prescription" in *Reiter*²¹ and *NCAA*²² represent neither an endorsement of the total welfare approach thought to have been urged by the phrase's creator, Judge Bork, nor a reference to the current understanding of the phrase as consumers' surplus. *Reiter* simply upheld a ruling authorizing consumers to sue to recover overcharges. *NCAA* condemned restrictions on price and out-

The question today is

what the standard

should be in assessing

the economic

consequences of a

practice or transaction.

¹⁵ See, e.g., Gordon Spivack, *The Chicago School Approach to Single Firm Exercises of Monopoly Power: A Response*, 52 ANTITRUST L.J. 651, 653 (1983); Harlan M. Blake & William K. Jones, *In Defense of Antitrust*, 65 COLUM. L. REV. 377 (1965); Robert Pitofsky, *The Political Content of Antitrust*, 127 U. PA. L. REV. 1051 (1979); Louis B. Schwartz, "Justice" and Other Non-Economic Goals of Antitrust, 127 U. PA. L. REV. 1076 (1979). More recently, Professor Hovenkamp demonstrated that a primary purpose of the Sherman Act was the protection of rivals—such as the small oil companies attacked and then acquired by Standard Oil. See Herbert Hovenkamp, *Antitrust's Protected Classes*, 88 MICH. L. REV. 1, 24–30 (1989). Cf. John S. McGee, *Predatory Price Cutting: The Standard Oil (N.J.) Case*, 1 J.L. & ECON. 137 (1958).

¹⁶ See, e.g., *Interface Group v. Mass. Port Auth.*, 816 F.2d 9, 10 (1st Cir. 1987) (Breyer, J.) ("'Anticompetitive' . . . refers . . . to actions that harm the competitive process, a process that aims to bring consumers the benefits of lower prices, better products, and more efficient production methods."); Werden, *supra* note 7. Werden identifies the goals of antitrust as those set forth famously in *Northern Pacific Railway*, and argues that the competitive process standard is the best means for achieving those goals. The passage from *Northern Pacific Railway* states: "The Sherman Act was designed to be a comprehensive charter of economic liberty aimed at preserving free and unfettered competition as the rule of trade. It rests on the premise that the unrestrained interaction of competitive forces will yield the best allocation of our economic resources, the lowest prices, the highest quality, and the greatest material progress, while at the same time providing an environment conducive to the preservation of our democratic political and social institutions." *Northern Pac. Ry. Co. v. United States*, 356 U.S. 1, 4 (1958).

¹⁷ See, e.g., Gordon Spivack, *Monopolization Under Sherman Act, Section 2*, 50 ANTITRUST L.J. 285, 304–07 (1982); Robert Lande, *Wealth Transfers as the Original and Primary Concern of Antitrust: The Efficiency Interpretation Challenged*, 34 HASTINGS L.J. 65 (1982). These articles, and many others (e.g., the articles cited *supra* note 15), demonstrate the inaccuracy of Judge Bork's argument that Congress's original intent was solely to maximize economic efficiency. BORK, *supra* note 1, ch. 2.

¹⁸ *Continental T.V., Inc. v. GTE Sylvania Inc.*, 433 U.S. 36 (1977).

¹⁹ *Brunswick Corp. v. Pueblo Bowl-O-Mat, Inc.*, 429 U.S. 477 (1977).

²⁰ Werden, *supra* note 7, at 737–43.

²¹ *Reiter v. Sonotone Corp.*, 442 U.S. 330, 343 (1979).

²² *NCAA v. Bd. of Regents*, 468 U.S. 85, 107 (1984).

put that would be prohibited under any standard. The Court has never addressed the standard in a context where it truly mattered; and so, as a matter of general jurisprudence, the welfare standard question must still be viewed as open.

In examining the five potential alternative standards, the analysis must include, not only whether we view a standard as doctrinally correct, but also its administrability. A standard that is “right” in the abstract has very little value if it cannot be applied practicably in court, in agency reviews, or, most importantly, in counseling clients.

Of the five alternatives, one of the easier ones to discard is the multiple goals concept. Although consistent with the original congressional intent of the law and prevalent for a long time, the courts abandoned it almost 40 years ago—and there has been no serious effort to reinstate it. That is so, at least in part, because it is often internally inconsistent and unadministrable.²³ For example, suppose a group of small dealers agrees to fix and raise prices. The enhanced profits would help “preserve” these small businesses, but would not disperse economic power or promote fairness in economic dealings. And consumers would certainly be harmed. Decision makers would be at a loss in deciding which of these conflicting goals has priority over others.

The “consumer choice” standard does not fare much better. Virtually every merger involving competing products will entail the exit or change of one or more products. That reduction in “choice,” in fact, is often the very source of the economic efficiencies that render so many mergers beneficial. As Judge Douglas Ginsburg and FTC Commissioner Joshua Wright have explained: “The flaw in this approach is that both economic theory and empirical evidence are replete with examples of business conduct that simultaneously reduces choice and increases welfare in the form of lower prices, increased innovation, or higher quality products and services.”²⁴ Another flaw is that the standard is necessarily arbitrary. Is a reduction in choices from 100 to 99 unreasonable? Or five to four? There is no objective way to tell.

In contrast, the “total” welfare standard has many adherents and much to commend it. Its premise is the prohibition of only those practices that reduce the wealth of society as a whole—which certainly sounds laudable. In terms of competition policy, however, that strength can also be a weakness. Professor Salop has provided an example that demonstrates the point: a merger (or conduct) that reduces the defendant’s costs, resulting in lower prices to consumers—but that also drives some rival producers out of business as a result.²⁵ If the harm to the rivals results in a loss of aggregate producer surplus that exceeds the gain to consumers, the merger would not be allowed. Similarly, the total welfare standard, rigorously applied, would condemn vertical restraints that lower prices to consumers if the loss to rivals is greater. These outcomes, of course, are contrary to longstanding precedent holding that antitrust protects competition, not competitors.²⁶ And not even the proponents of a total welfare standard defend these results. Analysis of these and similar examples demonstrates that what we are really interested in is the process of competition, not textbook economic welfare as a whole.

[W]hat we are really interested in is the process of competition, not textbook economic welfare as a whole.

²³ This point was ably demonstrated by Judge Bork. BORK, *supra* note 1, chs. 2–3.

²⁴ Joshua D. Wright & Douglas H. Ginsburg, *The Goals of Antitrust: Welfare Trumps Choice*, 81 *FORDHAM L. REV.* 2405, 2411 (2012). It is also worth noting that consumer choice is not a “welfare” standard: it does not purport to measure surplus in any respect, producer, consumer, general, or otherwise.

²⁵ See Salop, *supra* note 10, at 343.

²⁶ *E.g.*, *Brunswick*, 429 U.S. at 488.

Probably the most widely favored standard today is the consumer welfare standard. It is commonplace to speak of antitrust as focused on consumer welfare, and to require claimants to make a demonstration of consumer harm. The Supreme Court said in *ARCO* that “[l]ow prices benefit consumers regardless of how those prices are set, and so long as they are above predatory levels, they do not threaten competition.”²⁷ Recognizing this point, the lower courts and federal enforcement agencies today consistently apply what they perceive to be a consumer welfare standard.

A consumer welfare focus also has some practical advantages over total welfare. Perhaps the most significant of these is relative ease of measurement. Under a consumer welfare test, if a practice yields lower prices or higher output, then that generally resolves the matter without a need for further inquiry. Under a total welfare test, in contrast, it is necessary to quantify and net out consumer losses against producer gains—a process that can be especially difficult in a litigation context.²⁸

The consumer welfare standard loses some of its appeal, however, when it is pointed out that, technically, the sole focus of that standard is on consumer surplus (as economists define the term). There are some practices—although quite few in number and fewer still as actual occurrences—in which competition is harmed even where consumer surplus increases. One example is a consumer-buyers’ cartel. If the cartel restricts its purchases—a reduction in market output—such that prices decline, the consequences will typically include a deadweight (or allocative efficiency) loss and a wealth transfer from producers to consumers.²⁹ In such a case, consumer surplus will increase, but competition is harmed. We do not want buyers going around entering into naked agreements to fix prices even if consumer surplus increases as a result.³⁰ The Department of Justice, in fact, will prosecute these types of cases criminally.³¹

The consumer surplus standard becomes especially complicated in dealing with certain vertical restraints. If, for example, a seller’s resale price maintenance increases market output for the product, but buyers pay more whether they want the resulting dealer services or not, is that an antitrust offense? Focusing only on consumers’ surplus may be misleading (and difficult to calculate) in terms of the economic effect on those consumers who pay a higher price in instances where the product itself is unchanged. And what of the consumers who would not buy the product at all but for the services induced by the resale maintenance program?³² Under a consumer surplus regime, similar issues arise in evaluating price discrimination practices and metering ties.³³

²⁷ *Atl. Richfield Co. v. USA Petroleum Co.*, 495 U.S. 328, 340 (1990).

²⁸ Under a total welfare test, moreover, if a practice actually harms consumers with lower output, then, in the usual case, producer gains cannot come from economies of scale because output will be lower than before. As a result, the most common source of producer gains is not available. See Herbert Hovenkamp, *Consumer Welfare in Competition and Intellectual Property Law*, 9 COMPETITION POL’Y INT’L, Autumn 2013, at 53, 56.

²⁹ *E.g.*, Salop, *supra* note 10, at 342. For a discussion of monopsony generally, see Jonathan M. Jacobson, *Monopsony 2013: Still Not Truly Symmetric*, ANTITRUST SOURCE (Dec. 2013), http://www.americanbar.org/content/dam/aba/publishing/antitrust_source/dec13_jacobson_12_18f.authcheckdam.pdf.

³⁰ See *Knevelbaard Dairies v. Kraft Foods, Inc.*, 232 F.3d 979, 988 (9th Cir. 2000) (injury to buyers sufficient even if lower prices to consumers result); see also Gregory J. Werden, *Monopsony and the Sherman Act: Consumer Welfare in a New Light*, 74 ANTITRUST L.J. 707 (2007).

³¹ *E.g.*, *United States v. Seville Indus. Mach. Corp.*, 696 F. Supp. 986 (D.N.J. 1988).

³² See generally *Leegin Creative Leather Prods. v. PSKS, Inc.*, 551 U.S. 877 (2007).

³³ A classic example of a metering tie is *IBM Corp. v. United States*, 298 U.S. 131 (1936).

The one standard that seems to defeat each of these criticisms is the competitive process standard—unadorned by any particular welfare requirement.

The one standard that seems to defeat each of these criticisms is the competitive process standard—unadorned by any particular welfare requirement. Under that approach, practices or transactions that impede the competitive process from working effectively fall into the prohibited category. This will include practices that restrict market output or that exclude rivals on bases other than efficiency while enhancing the defendant's market power. As an example of a prohibited practice, consider the *Professional Engineers* case.³⁴ The defendants implemented a professional code that banned competitive bidding on the grounds that unabated price reductions could reduce building quality and safety. The Court found the conduct illegal and rejected the defense as “nothing less than a frontal assault on the basic policy of the Sherman Act” because it was based on the premise that competition itself was harmful.³⁵

Conversely, consider the *Discon* case.³⁶ The conduct at issue involved alleged regulatory deception that allowed the defendant to raise prices but with negative impacts on at most a single rival, rather than the competitive process. Because the higher prices were the result of gaming the regulatory system, and not “from a less competitive market,” the claim was rejected.³⁷ Consumers were harmed on those facts, but not by a violation of the antitrust laws. Similarly, tie-in or bundling arrangements that may cause buyers to pay more—like cable television program bundles—are generally not prohibited absent some reduction in competition in the tied product market. Without such an effect, one cannot say that the *competitive process* has been harmed.³⁸ *Professional Engineers* and *Discon*, taken together, show that conduct that does not implicate the process of competition is not outlawed even if consumers are harmed, while conduct that impairs the competitive process is subject to condemnation absent a justification that, if proven, would demonstrate that the competitive process has actually been enhanced.

Gregory Werden's article³⁹ makes a convincing argument that the competitive process standard is the one standard that is truly consistent with both the Supreme Court's case law over many years and the economic underpinnings of modern antitrust. This standard, however, does not provide the complete answer we are seeking. Since *Sylvania* in 1977, proof of economic harm has been essential to any antitrust case,⁴⁰ but saying that a practice interferes with the competitive process does not tell us what kind of economic harm is required. Something more is needed—an understanding of the *type of anticompetitive effect* the antitrust laws are designed to prevent. Without that understanding, the exercise can be circular.

Output and the Competitive Process

The multiple goals and consumer choice standards are unworkable. The total welfare standard can be effective, but generates obviously incorrect results in a number of instances. The consumer welfare standard comes very close, but also misses the mark in enough contexts to call its utility into doubt as a universal answer. The one that emerges best is the competitive process stan-

³⁴ Nat'l Soc'y of Prof'l Eng'rs v. United States, 435 U.S. 679 (1978).

³⁵ *Id.* at 695. This and many other examples are discussed in some detail in Gregory Werden's article articulating a competitive process standard. Werden, *supra* note 7.

³⁶ NYNEX Corp. v. Discon, Inc., 525 U.S. 128 (1998).

³⁷ *Id.* at 136.

³⁸ Werden, *supra* note 7, at 758 & n.285; *see generally* Brantley v. NBC Universal, Inc., 675 F.3d 1192 (9th Cir. 2012).

³⁹ Werden, *supra* note 7.

⁴⁰ *See* Heyer, *supra* note 5, at 2.

dard, but it needs an added factor to make it more effective in defining anticompetitive effects. The added factor recommended here is market output.⁴¹

A focus on output is consistent with antitrust policy and practice from the beginning. Section 1 of the Sherman Act speaks of “restraint of trade.”⁴² The Supreme Court’s earliest decisions on the merits—*Trans Missouri* and *Joint Traffic*—condemned regimes that increased price and decreased output,⁴³ while the *Standard Oil* decision announcing the rule of reason specified the “limitation of production” as one of the key evils the law was designed to prevent.⁴⁴

Judge Bork’s *Antitrust Paradox* makes the case that an antitrust policy designed to prevent agreements and practices that reduce market output is consistent with the purpose of the law, provides an administrable mechanism to guide enforcement, enhances economic efficiency, and furthers the competitive process.⁴⁵ Professor Hovenkamp’s newly added subchapter on the subject, similarly, expresses the view that antitrust’s “overall goal is markets that maximize output, whether measured by quantity or quality.”⁴⁶

A focus on output has many virtues. Decreased output generally means higher prices. An output reduction will also typically lead to a deadweight welfare loss and associated diminution of allocative efficiency—the *bêtes noires* of the total welfare approach. Reduced output is also typically associated with a transfer of wealth from consumers to sellers, and an associated reduction in consumer surplus. So an output measure is largely consistent with both the total and consumer welfare paradigms. And conduct that causes a reduction in market output will often be connected to some interference with the competitive process.

None of this is to suggest that output is a panacea. Output can be very hard to measure precisely, and the measurement must factor in differences in quality. The key is the impact on net output, taking into consideration the many separate facets consumers value. But while measuring all the relevant attributes may be achievable in some cases, it will be difficult or impossible in many others. One of the most important of these facets is innovation, which can be especially difficult to quantify. Yet, innovation effects must be taken into account because innovation is the source of much of the gains accruing to society over time. A further complication is that the analysis must focus on the very-difficult-to-measure output that would have been produced “but for” the restraint in issue, comparing it with the output that was produced with the restraint in place—a particularly challenging task in rapidly growing industries.

In addition, output alone cannot be the test. Much conduct, such as simply going out of business, “reduces output,” but not in any way that implicates antitrust policy. And some conduct that increases output, such as predatory pricing, is appropriately prohibited when there is reason to believe that the longer-run effects will be negative. Output must be used in connection with the overall competitive process standard to determine whether competition has been harmed.

⁴¹ Adding an output gloss to the consumer welfare standard would have a similar effect. The reason for favoring competitive process/output is that it is more consistent with the case law, especially cases like *Discon*, where consumer surplus was diminished but the competitive process (at least in the Court’s eyes) was not harmed.

⁴² 15 U.S.C. § 1.

⁴³ *United States v. Trans-Missouri Freight Ass’n*, 166 U.S. 290 (1897); *United States v. Joint Traffic Ass’n*, 171 U.S. 505 (1898).

⁴⁴ *Standard Oil Co. v. United States*, 221 U.S. 1, 52 (1911).

⁴⁵ BORK, *supra* note 1, at 35 & *passim*.

⁴⁶ AREEDA & HOVENKAMP, *supra* note 3, ¶ 114a. Professor Hovenkamp favors a consumer welfare approach, but with some exceptions and with a focus on output. *See id.* ¶¶ 114b, 114e. Overall, his approach seems quite similar to the one suggested here.

*[C]onduct that reduces
output should be
condemned if the
competitive process
has been impaired.*

These complications should not deter the use of an output test under a competitive process regime. When litigating per se cases, a negative effect on output is presumed. In rule of reasons cases, proof of an anticompetitive effect is part of the plaintiff's initial burden, and an adverse effect on output should be an implicit part of that burden. Most of the time, there will be no need to measure output. In cases involving purely vertical territorial or customer restraints, the possibility of a reduction in output is so remote that a rule of virtual per se legality would make sense.⁴⁷ With mergers, the traditional tools used under the Merger Guidelines serve as a useful proxy for output, although efficiencies yielding non-price benefits may be entitled to greater weight under an output standard than under the Guidelines' focus on price effects.⁴⁸ Exclusionary conduct cases, both vertical and unilateral, will remain hard, as they so often are, but it is in those cases that an output focus will be most valuable in distinguishing the harmful from the benign.⁴⁹ Conduct that, overall, does not decrease market output should be upheld, and conduct that reduces output should be condemned if the competitive process has been impaired.

If antitrust courts and enforcers can direct their primary focus to conduct that impairs the competitive process, and rely on output effects to determine close calls on whether that process is truly being harmed, we should reach the right result in all but the most exceptional cases. Until something better comes along, this seems to be the best way to go. ●

⁴⁷ Cf. Richard A. Posner, *The Next Step in the Antitrust Treatment of Restricted Distribution: Per Se Legality*, 48 U. CHI. L. REV. 6 (1981).

⁴⁸ Horizontal Merger Guidelines, *supra* note 9, § 10.

⁴⁹ See Jonathan M. Jacobson & Scott A. Sher, "No Economic Sense" Makes No Sense for Exclusive Dealing, 73 ANTITRUST L.J. 779 (2006).

Innovators, Implementers, and Two-sided Hold-up

Luke Froeb and Mikhael Shor

In February of 2015, the U.S. Department of Justice issued a Business Review Letter regarding proposed revisions to the Institute of Electrical and Electronics Engineers (IEEE) Patent Policy.¹ This policy focuses on commitments by firms contributing patented technology to standards promulgated by the IEEE to license their technology on reasonable and non-discriminatory terms. Two main proposals seem designed to shift bargaining rents toward implementers and away from the developers of technology: (1) curtailing injunctive relief and (2) basing adjudicated royalties on the “smallest salable Compliant Implementation.” While purportedly designed to solve the “hold-up problem” faced by implementers of patented technologies, we are concerned that such proposals are likely instead to reduce innovation. This article does not focus on the specific elements of the IEEE revisions but considers these types of hotly contested proposals generally as they are being thought about today.

■
Luke Froeb is the
William C. Oehmig
Associate Professor
of Free Enterprise and
Entrepreneurship at
the Owen Graduate
School of Management
at Vanderbilt University
and a former Director
of the Bureau of
Economics at the Federal
Trade Commission.

Mikhael Shor is an
Associate Professor
of Economics at
the University of
Connecticut. The
authors have previously
consulted for Ericsson
on this issue. This article
expresses their own
views, and does not
necessarily reflect those
of Ericsson.

Background

A patent becomes a standard-essential patent (SEP) when compliance with an adopted standard necessitates use of the patent’s underlying technology. Out of concern that holders of SEPs could leverage the market power derived from essentiality, standard-setting organizations such as the IEEE and the European Telecommunications Standards Institution (ETSI) generally require holders of SEPs to make available a license on “reasonable and non-discriminatory” (RAND) terms.² This aims to reduce so-called patent “hold-up.”³ Patent hold-up may occur when the owner of a SEP makes demands for royalties after a potential implementer has invested in a product incorporating the standard. These royalties may be higher than what would have been negotiated before the implementer irrecoverably invested. If implementers foresee this possibility, they may under-invest or not invest at all.

¹ Business Review Letter from Hon. Renata B. Hesse, Acting Assistant Att’y Gen., U.S. Dep’t of Justice, to Michael A. Lindsay, Esq., Dorsey & Whitney, L.L.P. (Feb. 2, 2015) [hereinafter IEEE Business Review Letter], available at <http://www.justice.gov/atr/public/busreview/311470.htm>.

² In some cases, “fair” is added to RAND (FRAND). ETSI Rules of Procedure art. 6.1 (Nov. 26, 2008). Judge Richard Posner argues that “the word ‘fair’ adds nothing to ‘reasonable’ and ‘nondiscriminatory.’” *Apple, Inc. v. Motorola, Inc.*, 869 F. Supp. 2d 901, 912 (N.D. Ill. 2012). Others contend that there is a distinction between RAND and FRAND and that RAND “focuses on ensuring economic efficiency issues, leaving aside fairness considerations.” Pierre Dehez & Sophie Poukens, *The Shapley Value as a Guide to FRAND Licensing Agreements*, 10 REV. L. & ECON. 265, 267 (2014).

³ See IEEE Business Review Letter, *supra* note 1 (“[T]he Update’s provisions also may further help to mitigate hold up . . .”). What economists call “hold-up” (or “opportunism”) is defined generally in OLIVER WILLIAMSON, *THE ECONOMIC INSTITUTIONS OF CAPITALISM* (1985). For an application to patents, see, for example, Mark Lemley & Carl Shapiro, *Patent Holdup and Royalty Stacking*, 85 TEX. L. REV. 1991 (2007).

To mitigate potential hold-up, some parties suggest that RAND should be determined by “postulating a hypothetical negotiation”⁴ between innovators and implementers at the point immediately prior to the patent being incorporated into the standard.⁵

In addition to requiring a RAND commitment by holders of SEPs, two further measures that purport to reduce patent hold-up have gained favor among some courts and academics and are reflected in the IEEE revisions. First is the proposal to adjudicate royalties based primarily on the “smallest salable” patent-practicing unit (or “smallest component”) rather than the patent’s contribution to the entire market value of the product.⁶ This proposal stems from a fear that basing royalties on the full value of the implementing technology can lead to excessive payments by implementers. Second is the proposal to curtail injunctive relief for holders of SEPs. Proponents of this proposal argue that the threat of injunction can allow innovators to extract supra-RAND royalties when it is costly for implementers to switch to alternate technologies.

Two-sided Hold-up

Concerns expressed regarding the potential for patent hold-up are generally one-sided, focusing on the incentives of the implementers of the technology while generally ignoring the incentives of innovators to create the technology in the first place.⁷ Such analysis takes the level of innovation as given, so that the only possible harm to competition is through the implementer’s decision to invest. The underlying incentive problem, however, is two-sided because an innovator’s incentives to engage in significant R&D may also be distorted by well-intentioned actions taken to correct the potential hold-up problem.

Once we consider not just the effects of the proposed revisions on prospective implementers of technology, but also on innovators, the problem becomes recognizable as one of two-sided hold-up.⁸ Just as implementers invest before knowing what end-product demand will be, so too must innovators invest before knowing whether an innovation will be implemented.

Our analysis draws from an economic model of two-sided hold-up by Ganglmair, Froeb, and Werden.⁹ The model’s primary message is that the innovator’s and the implementer’s hold-up

⁴ *Minco, Inc. v. Combustion Eng’g, Inc.*, 95 F.3d 1109, 1119 (Fed. Cir. 1996).

⁵ Letter from Michael A. Lindsay, Esq., Dorsey & Whitney, L.L.P., to Hon. William J. Baer, Assistant Att’y Gen., U.S. Dep’t of Justice at 15 (Sept. 30, 2014) [hereinafter IEEE Business Review Request], available at <http://www.justice.gov/atr/public/busreview/request-letters/311483.pdf> (defining reasonable rate as “excluding the value, if any, resulting from the inclusion of that Essential Patent Claim’s technology in the IEEE standard.”).

⁶ “The value that the Essential Patent Claim contributes to the smallest saleable Compliant Implementation that practices that claim, in light of the value contributed by all Essential Patent Claims for the same IEEE Standard practiced in that Compliant Implementation.” *Id.* at 16. Our focus is not on the specific implementation of the rule (whether smallest component is defined economically, technologically, or in some combination), but on its general implication potentially limiting the scope of bargaining to only some of the patent’s value to the final product.

⁷ “Simply put, the FTC is concerned that a patent holder may . . . ‘hold-up’ or demand higher royalties or other more costly licensing terms after the standard is implemented than could have been obtained before its IP was included in the standard.” Prepared Statement of the Federal Trade Commission Before the United States Senate Committee on the Judiciary Concerning “Oversight of the Impact on Competition of Exclusion Orders to Enforce Standard-Essential Patents” (July 11, 2012) available at http://www.ftc.gov/sites/default/files/documents/public_statements/prepared-statement-federal-trade-commission-concerning-oversight-impact-competition-exclusion-orders/120711-standardpatents.pdf.

⁸ The “one lesson of economics” is that to evaluate policy you have to consider all of its effects, both immediate and longer term, and not just for one party but for all parties. HENRY HAZLITT, *ECONOMICS IN ONE LESSON* (1946).

⁹ Bernhard Ganglmair, Luke M. Froeb & Gregory J. Werden, *Patent Hold Up and Antitrust: How a Well-Intentioned Rule Could Retard Innovation*, 60 J. INDUS. ECON. 249 (2012).

problems are not directly comparable. It is possible for negotiations to occur prior to the implementer's investment in the standard, but negotiations always occur after the innovator's investment in R&D is sunk.

If innovators can extract additional rents after implementers invest in a standard, implementers foreseeing this possibility may fail to adopt or adequately invest in the standard in the first place. However, this underinvestment is not the result of royalty payments that are too high or too low, but of the fact that any bargaining that carries the potential for hold-up occurs after the implementer's investments are sunk.¹⁰ Bargaining that occurs *ex ante*, before the implementer sinks its investment in the standard, can lead to high or low royalty payments (depending on bargaining power), but in all cases can induce optimal investment by the implementer.¹¹ This is because both the implementer and the innovator benefit from the implementer's investments, as they increase the value of the product and thus the surplus over which the parties bargain. Since both parties prefer to increase the surplus over which they are bargaining, the negotiation is likely to provide an incentive for the implementer to invest optimally. A number of contractual (and litigation) solutions are available to ameliorate the implementer's hold-up problem.¹² Conversely, we believe that the innovator's hold-up problem is more difficult to overcome. The challenge is to provide sufficient incentive for the innovator to invest in R&D in the first place.

The challenge is to provide sufficient incentive for the innovator to invest in R&D in the first place.

The reason that innovators may have insufficient incentive to invest parallels the logic of the implementer's hold-up problem. Bargaining after the implementer invests leads to inefficient investment by the implementer. Similarly, bargaining after the innovator invests runs the risk of significantly altering investment away from what is socially optimal. Yet, what courts call an *ex ante* "hypothetical negotiation" is actually *ex interim*. It occurs before the implementer makes any irrevocable investments in the standard, but after the innovator has irrevocably invested in (and borne the risks of) research, development, and patenting of the innovation.¹³ Innovation costs are treated as "sunk" in the RAND *ex ante* model of hypothetical bargaining, but an innovator's investments depend crucially on the expected returns from its patent and the outcome of this bargaining. Because of the *ex interim* nature of bargaining, economic theory suggests that shifts in relative bargaining power have bigger effects on innovators' investments than on implementers' investments.¹⁴

Both the smallest component rule and curtailing of injunctive relief serve to shift bargaining power and profits from innovators to implementers. This shift weakens the value of patents and can significantly reduce the incentive to innovate.

¹⁰ *Id.* at 260–61.

¹¹ *Id.* at 256, 264.

¹² Several contractual forms, including RAND commitments, can resolve the implementer's incentive problem. For example, an option-to-license contract specifies in advance a maximum fee (whether a lump sum or an agreed royalty rate) that will be paid by the implementer if it uses the patented technology. Such a commitment made *ex ante* also provides the implementer with optimal incentives, but may offer the innovator a greater incentive to innovate than RAND does. *Id.* at 264 (Proposition 2). Additionally, an innovator's concern with its reputation and with future repercussions of any opportunistic behavior can ameliorate the hold-up problem. Damien Geradin, Anne Layne-Farrar & A. Jorge Padilla, *The Complements Problem within Standard Setting: Assessing the Evidence on Royalty Stacking*, 14 B.U. J. Sci. & Tech. L. 144 (2008).

¹³ Investments in R&D (as well as the innovator's investments in developing and marketing the resulting technologies and making specific investments toward incorporating the patented technology into the standard) are already sunk at the *ex interim* stage of negotiation, so, in our view, RAND incorporates a systematic "hold-up" of these investments.

¹⁴ Ganglmair, Froeb & Werden, *supra* note 9, at 260–61 (demonstrating that *ex interim* interpretation of RAND leads to the same optimal investment by the implementer for any level of bargaining power).

The Smallest Component Rule

The smallest component rule generally restricts the innovator to bargain only over a component for which it is “the ‘basis for customer demand.’”¹⁵ Technically, the smallest component rule does not constrain what parties could consider when negotiating but only changes what a court would consider if RAND terms were litigated. However, “all bargaining is necessarily done in the shadow of the law.”¹⁶ If the adjudicated RAND royalty is expected to rely on only the smallest component, then the implementer is less likely to accept royalty payments based on the entire market value as it can always (after a judicial proceeding) receive the smaller adjudicated royalty. The manufacturer is less likely to realize some or any of the increased demand generated by the patent beyond the value of the smallest component. This implies that the innovator will not consider this value when deciding whether or not to innovate.

Two conditions encourage optimal innovation. First, implementers and innovators should bargain over the entire additional value created by the patent.¹⁷ Second, innovators should have enough bargaining power to realize a significant enough portion of that surplus to make R&D worthwhile.¹⁸ The smallest component rule, by reducing the potential bargaining range and shifting bargaining power to implementers, fails both conditions.

The DOJ commented that a goal of RAND commitments is “assuring implementers that they will not have to pay any hold-up value connected with the standardization process.”¹⁹ We agree but, in our view, this is accomplished primarily not through any specific division of surplus but through ensuring the timing of negotiation (actual and hypothetical) is prior to the standardization and implementation.²⁰ Conversely, as negotiation cannot occur prior to innovation, the shifting of rents can adversely affect the innovator’s incentives. To put the economic argument simply, if a patent provides even a dollar of incremental value to the implementer beyond what is reflected in the smallest component, then that is a dollar that economic efficiency would require the innovator to appropriate in part. As the smallest component rule runs contrary to this, we expect its effect to be socially harmful—i.e., less innovation without any compensating increase in implementers’ investments.²¹

Royalty Stacking. Aside from focusing royalties on the smallest component, the IEEE also proposes to set royalties “in light of the [smallest component’s] value contributed by all Essential Patent Claims for the same IEEE Standard.”²² The DOJ notes that proper apportionment of value

¹⁵ *Rite-Hite Corp. v. Kelley Co.*, 56 F.3d 1538, 1549 (Fed. Cir. 1995) (emphasis added). Some variations in language have also been proposed, allowing for the patent to constitute a “predominant” or “substantial” basis for demand. *Lucent Technologies, Inc. v. Gateway, Inc.*, 580 F.3d 1301, 1337 (Fed. Cir. 2009) (“the basis—or even a substantial basis—of the consumer demand”); Patent Reform Act of 2009, H.R. 1260, 111th Cong. § 284(c)(1)(A) (2009) (“[C]ontribution over the prior art is the predominant basis for market demand.”). Our discussion is meant to be illustrative of the general approach, and our conclusions do not depend upon these particulars.

¹⁶ James Ratliff & Daniel L. Rubinfeld, *The Use and Threat of Injunctions in the RAND Context*, 9 J. COMPETITION L. & ECON. 1, 2 (2013).

¹⁷ Ganglmair, Froeb & Werden, *supra* note 9, at 256.

¹⁸ *Id.* at 256–57, 261–62.

¹⁹ IEEE Business Review Letter, *supra* note 1.

²⁰ This is precisely the goal of the IEEE’s mandatory factor in determining a reasonable rate that excludes the value of standardization. It is unclear to us how the smallest component rule further contributes to reducing patent hold-up.

²¹ Some may argue that the smallest component rule might reduce the likelihood of hold-up of implementers. While we do not see why this would occur (aside from arbitrarily shifting rents to implementers), the smallest component rule will certainly reduce returns to innovators and hence incentives for innovation.

²² IEEE Business Review Request, *supra* note 1, at 12.

in the context of a smallest component rule may ameliorate royalty stacking.²³ Royalty stacking is a fear that the sum of royalties across multiple SEP owners will make the technology economically infeasible to market. It is especially acute when the patents are complementary, so that the product's total value is not merely a sum of the incremental values of each of its patents.

This complementarity problem is neither new nor unique to patents. It was first recognized nearly 200 years ago,²⁴ and economists have made significant progress in addressing it in the interim. Effectively, royalty stacking occurs if negotiated royalty rates are set without appropriate thought given to the rates of other, complementary patents. Real negotiations rarely proceed down such a naïve path; no hypothetical negotiation should either. Of course one should set royalties in light of other patents, but carefully. The smallest component rule can already reduce incentives to innovate by limiting the amount over which an innovator bargains; if a proposed solution to royalty stacking further reduces this amount, then it only exacerbates the problem.

Royalty stacking should be avoided in litigation not by artificially reducing what is bargained over but by more realistically thinking how *ex ante* bargaining would occur.²⁵ A number of economic approaches exist to determine the appropriate sharing of surplus between the multiple innovators and the implementer, including both strategic equilibrium-based theories from non-cooperative game theory²⁶ and bargaining-based theories from cooperative game theory.²⁷ These models can provide likely royalties from an implementer's independent but simultaneous negotiations with several innovators of complementary patents. These approaches, if adopted, would simply extend what courts already consider in bilateral hypothetical negotiations to multilateral (or simultaneous bilateral) settings, and also allows for asymmetries, bargaining power, and other real-world concerns.

Conversely, changing the royalty base to the smallest component by itself does not solve the allocation problem among complementary patents. First, when the smallest salable patent-practicing unit also contains multiple patented features or a mix of patented and unpatented features, it inherits the exact same economic challenges of the end-user product. Second, rather than solving the allocation problem, it simply sidesteps it by assuming that economic rents above those in the smallest component should flow primarily to the implementer of the end product. This "resolves" the allocation problem by allocating none of the surplus to innovators when no single patent can lay major claim to it.

²³ *Id.* at 13 ("Regarding the second recommended factor, appropriately apportioning the value of all essential patent claims in an IEEE standard addresses royalty stacking, which may hamper implementation of a standard.").

²⁴ AUGUSTIN COURNOT, RESEARCHES INTO THE MATHEMATICAL PRINCIPLES OF THE THEORY OF WEALTH 99 (Nathaniel T. Bacon trans., 1838) ("Ordinarily . . . several raw materials are generally brought together in the manufacture of each of these products Hence it is necessary to inquire according to what laws the profits, which are made by all of the producers as a whole, are distributed among the individuals in consequence of the law of consumption").

²⁵ Here, artificial reduction refers both to the use of the smallest component rule and to potential naïve solutions to royalty stacking, such as limiting the bargaining range to $1/N$ of the total component value, where N is the number of patents. "There are at least 92 entities that own 802.11 SEPs. If each of these 92 entities sought royalties similar to Motorola's request of 1.15% to 1.73% of the end-product price, the aggregate royalty to implement the 802.11 Standard, which is only one feature of the Xbox product, would exceed the total product price." *Microsoft Corp. v. Motorola, Inc.*, Case No. C10-1823 (W.D. Wash. Apr. 25, 2013).

²⁶ Daniel Quint, *Pooling with Essential and Nonessential Patents*, 6 AM. ECON. J.: MICROECONOMICS 23 (2014); Carl Shapiro, *Navigating the Patent Thicket: Cross Licenses, Patent Pools and Standard Setting*, in 1 INNOVATION POLICY AND THE ECONOMY 119 (Adam Jaffe, Joshua Lerner & Scott Stern eds., 2001).

²⁷ See Dehez & Poukens, *supra* note 2; Anne Layne-Farrar, A. Jorge Padilla, & Richard Schmalensee, *Pricing Patents for Licensing in Standard-Setting Organizations: Making Sense of FRAND Commitments*, 74 ANTITRUST L.J. 671 (2007).

The correct solution to royalty stacking concerns is not artificially limiting the base for each patent royalty, but is the proper allocation of patents to the product's entire market value.²⁸

Product Design. While the smallest component rule, in our view, solves neither two-sided hold-up nor royalty stacking, it does create perverse and inefficient incentives. Economic analysis of bargaining generally contemplates two steps. In the first step, firms work to maximize the net value of the final product by, for example, producing components at lowest cost and providing consumers with products tailored to their needs. This provides the firms with more surplus (consumer value minus cost) to divide among themselves. In the second step, firms split the resulting surplus through a bargaining process that accounts for the bargaining strength of each party,²⁹ the value that each contributes, the availability of alternatives, and market and other factors.

Generally, economists are less concerned with the second stage—how firms choose to split the surplus—except when its procedures unnecessarily affect the incentives in the first stage. The smallest component rule does precisely this by introducing patent royalty issues into product design. This shifts incentives of the parties away from maximizing total surplus and towards designing products tailored to exploiting (or “gaming”) the smallest component rules.

This would cause at least two distortions, each reducing the normal bargaining incentives to maximize total surplus. First, to the extent that the choice of royalty base is a function of the importance of the patented technology to each component of the device, an implementer may desire to minimize the possibility that the patent constitutes a “substantial” driver of demand. This creates an incentive for the implementer to diminish the patent's importance to the final product, perhaps by inefficiently reducing reliance on the patented technology or by introducing other (perhaps unnecessary) product features to dilute a SEP's relative importance.

Second, both the innovator and the implementer have incentives to design intermediate products not only in terms of their effectiveness and ease of integration but also with an eye on whether they (or some alternatives) would constitute the smallest salable component. These design considerations create potential tradeoffs between the economic concern of increasing surplus and the entirely artificial concerns of product design for the sole purpose of skirting or tailoring the product to exploit or game the bargaining rules. Any such concerns (other than surplus maximization) create inefficient distortions, increasing costs or reducing the value of the product. The reduction of surplus inherent in designing around the smallest component rule is likely to lead to underinvestment by innovators in research and development of SEP technologies.

Multiple Valued Uses. A single patent may find itself in many applications with varied uses and values. But what if the same “smallest component” is used in every application? If the smallest component rule requires the innovator to charge each the same royalty or makes differential pricing more difficult,³⁰ then it limits the innovator's ability to internalize the variety of valued uses of the innovation by precluding cases where different types of contracts would be arranged with different types of implementers. If this is the case, then the smallest component rule runs contrary to the RAND principle of providing a reasonable return on the innovator's investment and would likely lead to an under-provision of innovation.

²⁸ Although the IEEE rule calls for a consideration of other patents, it does not specify how this is to occur other than limiting the size of the royalty base.

²⁹ “The precise point royalty within the bargaining range should be informed by the relative bargaining power of the infringer and patent holder.” J. Gregory Sidak, *The Meaning of FRAND, Part I: Royalties*, 9 J. COMPETITION L. & ECON. 931, 939 (2013).

³⁰ For example, it may be difficult to monitor contracts with a component manufacturer when those contracts stipulate different prices for the same device depending on their ultimate use.

The correct solution to royalty stacking concerns is not artificially limiting the base for each patent royalty, but is the proper allocation of patents to the product's entire market value.

For example, imagine that the same component (incorporating the same essential patent) is used to help stabilize flight of both commercial airplanes and toy airplanes. Clearly, these industries are likely to have different values for the patent. By negotiating over a single royalty rate based on the component price, the innovator would either fail to realize the added value of its patent to commercial airlines, or (in the case that the component is targeted primary to the commercial airlines) would not realize the incremental market potential from the patent's use in toy airplanes. In either case, the innovator will not be negotiating over the entirety of the value it creates, leading to too little innovation.

Cognitive Biases and Jurors. One argument made in favor of the smallest component rule is rooted in a concern that jurors may have preconceived notions of “reasonable” percentage royalty rates. For concreteness, note that 10% of \$100 is mathematically identical to 1% of \$1,000 and to 0.1% of \$10,000, and thus the same royalties (in dollar terms) can be computed from any of these three bases with a corresponding adjustment to the royalty rate. However, if jurors see rates like 0.1% as “too small,” then (the argument goes), using the entire market value—which may be associated with small royalty rates—will lead to overpayment when juries reject those rates in favor of more “focal” or “cognitively” reasonable ones.

There are several reasons why the smallest component rule is not an appropriate solution to this perceived problem. First, adjusting the “base” so that it leads to desired royalties when multiplied by the jurors’ preconceived royalty rate is entirely circular and presupposes that one has an idea of what the right royalties are. Second, this is not a problem unique to the use of the entire market value as the royalty base. Too small of a “smallest component” could require a correspondingly large royalty rate which exceeds jurors’ focal points and thus would be revised downward, underpaying the innovator. Third, the proper way to overcome biases is to mitigate them through juror education, not to tailor jurisprudence to them. Research in psychology and behavioral economics suggests some possible solutions.³¹ Sometimes something as uncomplicated as reframing the decision (“\$1 for every \$1,000” instead of “zero point one percent”) is sufficient to ameliorate such biases. Simply, this is a lawyering, not a law, problem.³²

Injunctive Relief

The proposal to weaken injunctive relief rests on a fear of hold-up by the innovator when implementers face switching costs to another technology. The innovator can, in theory, use the threat of injunction to extract fees in excess of those that might have prevailed if bargaining had occurred *ex ante*. Based on this reasoning, some economists have argued for curtailing injunctive relief. Often, these arguments ignore or downplay the importance of an innovator’s incentives.

For example, in an influential paper cited in the *IEEE Business Review Letter*, Lemley and Shapiro argue that injunctive relief leads to overcharges and therefore should be disallowed to certain classes of patent holders.³³ However, these purported overcharges are with respect to a

³¹ See, for example, Dale J. Cohen, Jennifer M. Ferrell & Nathan Johnson, *What Very Small Numbers Mean*, 131 J. EXP. PSYCHOL. 424 (2002).

³² Some experiments involving jurors have found that “debiasing” strategies can help overcome at least one form of cognitive bias (“hindsight bias”). D. Jordan Lowe & Philip M.J. Reckers, *The Effects of Hindsight Bias on Jurors’ Evaluations of Auditor Decisions*, 25 DECISION SCI. 401 (2007).

³³ Lemley & Shapiro, *supra* note 3.

“natural benchmark” that reflects the innovator’s bargaining power, and not with respect to the royalties that encourage socially optimal innovation.³⁴

Conversely, when the incentives to innovate are properly considered, royalties in the absence of injunctions are often well below those that encourage efficient, socially optimal levels of innovation.³⁵ As long as contracting occurs prior to the implementer’s investment, the implementer’s incentive to invest is not a function of bargaining power. However, the degree of investment in innovation is a function of each party’s bargaining power.³⁶ Therefore, policies that further reduce the innovator’s bargaining power can result in under-provision of innovation or a refusal to participate in standards that call for such reductions in bargaining power.³⁷ A reduction in the availability of injunctive relief shifts bargaining power away from the innovator.

Prohibition of injunctions may or may not preclude any hold up of implementers, but will certainly reduce negotiated royalty rates. The resulting shift in bargaining power from innovators to implementers “reduces the payoff from R&D and makes some projects no longer worth pursuing.”³⁸

Injunctions are blunt tools that can negatively affect holders of related patents, non-infringing products, consumers, and manufacturers. The potential for efficiency loss and the importance of relative bargaining power speak to a careful balancing of effects on a case-by-case basis, as is already a cornerstone of jurisprudence relating to the granting of injunctions outside of SSOs.³⁹ Arbitrary restrictions or categorical bans are the opposite of balancing.

Access to Injunctive Relief Restores Ex-Ante Bargaining. One argument made in favor of curtailing injunctive relief is that it is inconsistent with RAND principles. Per this argument, the innovator has effectively relinquished the right to injunctive relief by agreeing to license the patent on RAND terms as a condition of its incorporation in the standard. Effectively, the argument is that the innovator has acknowledged that monetary damages are sufficient under RAND.

One often-overlooked aspect of the RAND commitment is that the innovator agrees to allow access to the patented technology to any and every firm, relinquishing one of the most important aspects of bargaining—the ability to refuse access. In *ex ante* bargaining before the patent became part of the standard, the innovator had the ability to make its technology unavailable to a potential implementer. In fact, *ex ante*, both the innovator and implementer of end-user products negotiate with full freedom to walk away. After the setting of a standard, bargaining becomes one-sided; the innovator must provide access, but the implementer has no obligation to market a product incorporating the standard or to pay a RAND fee.

³⁴ Mark A. Lemley & Carl Shapiro, *Patent Holdup and Royalty Stacking: Reply*, 85 TEX. L. REV. 2163, 2166 (2007). Other commentators also note that Lemley and Shapiro’s proposed remedy of removing injunctive relief “is not calculated to maximize any objective measure of welfare.” J. Gregory Sidak, *Holdup, Royalty Stacking, and the Presumption of Injunctive Relief for Patent Infringement: A Reply to Lemley and Shapiro*, 92 MINN. L. REV. 714, 747 (2008). The royalty rates in Lemley and Shapiro often fall below socially optimal rates. Einer Elhauge, *Do Patent Holdup and Royalty Stacking Lead to Systematically Excessive Royalties?*, 4 J. COMPETITION L. & ECON. 535 (2008).

³⁵ Ganglmair, Froeb & Werden, *supra* note 9, at 261 (Lemma 5).

³⁶ *Id.* at 260 (Lemma 4).

³⁷ Benjamin Chiao, Josh Lerner & Jean Tirole, *The Rules of Standard?Setting Organizations: An Empirical Analysis*, 38 RAND J. ECON. 905 (2007).

³⁸ Ganglmair, Froeb & Werden, *supra* note 9, at 265.

³⁹ In the context of standard setting, it is “unlikely that a patent holder bound by a RAND commitment, even one that does not address explicitly the availability of injunctive relief, can secure an injunction (in addition to monetary damages) in an infringement action.” Therefore, the DOJ argues that “in practice, it [IEEE proposed rule limiting injunctive relief] will not be significantly more restrictive than current U.S. case law.” IEEE Business Review Letter, *supra* note 1, at 9–10.

RAND purports to reflect the negotiation and relative bargaining power that would have occurred absent the existence of a standard. Yet, the RAND commitment prohibits the innovator from withdrawing access to its technology, eliminating a major source of the innovator's bargaining power. Injunctions provide for very specific circumstances under which an infringer can be denied the access to the technology. Therefore, the possibility of injunction can be seen as partially restorative of the bargaining power that the innovator had before the existence of the standard.⁴⁰

Access to Injunctive Relief Encourages Negotiation. The potential implementer chooses between negotiating license royalties now and paying damages later, if found infringing by a court of law. As RAND damages are generally conceived as being equal to the royalties that would have been negotiated, the implementer chooses between licensing—paying a fee now for certain—and infringing—paying roughly the same fee later, but only if the infringement is discovered and proven. The absence of injunctive relief, coupled with an irrevocable RAND commitment, can imply that infringement is cheaper than licensing. This is likely to lead to a constructive refusal to negotiate by implementers. What should be actual *ex-ante* negotiations would be replaced with hypothetical ones at trial.⁴¹ For the innovator, this implies expected returns on innovation that are below RAND, as they must account for the necessary litigation costs.⁴²

This is not merely an academic concern. For example, the International Trade Commission concluded that, in its protracted infringement of Samsung's patents, "Apple has no intention of paying Samsung any royalties until after the conclusion of litigation"⁴³ even though Samsung negotiated in good faith.⁴⁴ The Commission noted that this unwillingness to negotiate in good faith forces the innovator to undertake "expensive litigation" or forgo royalties entirely.⁴⁵

Conclusion

Concerns that implementers may overpay for royalty licenses, thereby discouraging investments made to implement the standard, are often put forth without giving equal consideration to the possibility of underpayment and its impact on innovators' incentives to create the technology underlying the standard.

The absence of injunctive relief, coupled with an irrevocable RAND commitment, can imply that infringement is cheaper than licensing.

⁴⁰ For example, FTC Commissioner Joshua D. Wright called the presumption that injunctions lead to supra-FRAND royalties "dubious," noting that the threat of injunction can be used to achieve RAND royalty rates. Joshua D. Wright, Comm'r, Fed. Trade Comm'n, SSOs, FRAND, and Antitrust: Lessons from the Economics of Incomplete Contracts, Remarks at Center for the Protection of Intellectual Property Inaugural Academic Conference: The Commercial Function of Patents in Today's Innovation Economy 29–30, George Mason University School of Law (Sept. 12, 2013), available at https://www.ftc.gov/sites/default/files/documents/public_statements/ssos-frand-and-antitrust-lessons-economics-incomplete-contracts/130912cpip.pdf. See also J. Gregory Sidak, *supra* note 29, at 1008 ("There is consequently no valid justification to assume that royalties negotiated under the threat of an injunction necessarily violate FRAND.").

⁴¹ Injunctions can serve to encourage negotiation. "If used judiciously by the courts, the threat of the imposition of an injunction can serve, when needed, to move patent disputes towards resolution." Ratliff & Rubinfeld, *supra* note 16, at 22.

⁴² "[T]he SEP holder cannot refuse a license nor seek an injunction and instead can at best use infringement litigation to obtain the same FRAND rates and terms it would have gotten through good-faith negotiations (had they occurred), but only at a higher cost due to litigation." Anne Layne-Farrar, *Moving Past the SEP RAND Obsession: Some Thoughts on the Economic Implications of Unilateral Commitments and the Complexities of Patent Licensing*, 21 GEO. MASON L. REV. 1093, 1104 (2014).

⁴³ Certain Electronic Devices, Including Wireless Communication Devices, Portable Music and Data Processing Devices, and Tablet Computers at 62 (Investigation No. 337-TA-794, U.S. Int'l Trade Comm., July 5, 2013 Commission Opinion, Public Version).

⁴⁴ *Id.*

⁴⁵ "Apple's position illustrates the potential problem of so-called reverse patent hold-up, a concern identified in many of the public comments received by the Commission. In reverse patent hold-up, an implementer utilizes declared-essential technology without compensation to the patent owner under the guise that the patent owner's offers to license were not fair or reasonable. The patent owner is therefore forced to defend its rights through expensive litigation." *Id.* at 63.

Our analysis shows that underpayment is a greater threat to economic efficiency. The reason is that innovators and implementers can and do bargain prior to the implementer's adoption of and investment in a standard and courts impose such "hypothetical" bargaining when determining royalties. Therefore, the implementer's efficient investment, which benefits both parties, is an essential part of these deliberations. However, bargaining does not occur prior to the innovator's investment in R&D. In fact, all such "*ex ante*" bargaining occurs *after* these investments are sunk.

Curtailling injunctive relief and basing royalties on the smallest salable component both pose the risk of under-rewarding innovators for their investments. This is likely to retard innovation, reduce incentives to participate in standards, and reduce economic welfare. ●

Reflections on the Debates Surrounding Standard-Essential Patents

Gregory K. Leonard

In the policy debates that surround standard-essential patents (SEPs), reasoned discussion is sometimes replaced by rhetoric that, while having some economic basis, is one-sided. This state of affairs may not be surprising given that the stakes are high and the battle lines are clearly drawn. However, economically sound policies toward SEPs must consider the complexities of the issues involved in the context of limited economic knowledge. In this article, I delve into the complexities related to some of the debated issues, illustrating why nuanced analysis is necessary to move the debates forward.

Limits on Economic Knowledge Concerning the Patent System and Innovation

From the point of view of social welfare, the central goal of the patent system is achieving the optimal amount of innovation, and thus the focus should be on designing the patent system (including the legal framework) and standard-setting organization (SSO) policies to provide the “right” incentives. However, the practical reality is that economists have only a limited understanding of the effects of strengthening or weakening the patent system and economists know even less about whether innovation currently is at, above, or below the optimal level. Thus, the frequent rhetorical references in the SEP debate to dire consequences for innovation from changing (or not changing) the patent system or SSO policies are not well supported by the state of economic knowledge.

The problems start with the definition of “innovation.” The definition of innovation (and, correspondingly, the definition of “innovative activities”) should not be limited to patent-protected inventions. Innovations may be “protected” through means other than patenting, such as exploiting first-mover advantage (or more generally implementing innovations faster than competitors) or by maintaining the innovation as a secret (i.e., trade secrets). Changing the patent system may change the incentives to patent, but this may or may not change the overall incentives to engage in innovative activities and thus the amount of innovation, properly defined. For example, a strengthening of the patent system that leads to higher royalties for patented inventions could also thereby decrease the returns to non-patented innovations used in the same products (e.g., product development innovations) and thus lower the incentives to engage in activities aimed at developing such innovations.

A second difficulty is that the predictions of theoretical economic models of innovative activity are highly dependent on their assumptions regarding, for example, whether innovation is cumulative, the number of potential areas for innovation, etc. Different models with different assumptions can give diametrically opposite predictions. For example, in a model of a single potential innovation, a patent system generally improves social welfare, while in a model with cumulative innovations, a patent system can harm social welfare. As Bronwyn Hall, a significant contributor to the

■ **Gregory K. Leonard** is a Partner, Edgeworth Economics LLC. The author has previously consulted for Cisco, Motorola, Google, HTC, Intel, Samsung, and Qualcomm, among others, and testified in *Ericsson v. D-Link* (on behalf of defendants), *In re Innovatio* (on behalf of defendants), *Microsoft v. Motorola* (on behalf of Motorola), and *Realtek v. LSI* (on behalf of Realtek). The author thanks Mario Lopez for helpful comments. The views expressed here are the author's own and do not necessarily reflect the views of any other person or company.

economic literature on patents, has stated, “As is often the case with models that admit the complexity of the world, the theoretical literature in this area produces ambiguous results with respect to incentives provided by patents.”¹ Thus, it is not possible, on the basis of a theoretical model, to make a general statement about the effects of some aspect of the patent system on the level of innovation or whether innovation is currently at the optimal level. At best, theoretical models alone only can identify the set of potential factors that may affect innovation and the pathways by which those potential factors would operate.

In light of the ambiguity of the theoretical models, empirical analysis is needed to sort out whether changing the balance between innovators and implementers would increase or decrease innovation. However, here again economists’ current state of knowledge is limited. Empirical analysis of this question is hampered by the limited extent to which the balance between innovators and implementers has varied over time. Such variation is needed to identify the effects of changes in the strength of the patent system on innovation. Empirical analysis is further hindered by difficulties in measuring the amount of innovative activity. A high-level summary of the empirical literature that has developed despite these challenges is that a stronger patent system encourages more patenting. However, the evidence for a stronger patent system leading to more innovation is, at best, weak and there is evidence suggesting that strengthening the patent system may sometimes reduce innovation. The case for a stronger patent system leading to more innovation is better supported for some industries (e.g., pharmaceuticals and medical devices) than others, but even for the former, conflicting empirical studies exist.² Importantly, scant empirical evidence exists regarding whether innovation is, or has been, at the optimal level.

Given the limits on knowledge, changes to the patent system should be undertaken with caution and with appreciation for the ambiguity of the potential effects on innovation and, more importantly, consumer welfare and economic efficiency.

Given the limits on knowledge, changes to the patent system should be undertaken with caution and with appreciation for the ambiguity of the potential effects on innovation and, more importantly, consumer welfare and economic efficiency.

Smallest Salable Practicing Unit

The prospect of litigation, including the costs of litigation and the potential outcomes of litigation, influences negotiated patent licenses, even in those cases where an agreement is reached without recourse to litigation.³ No rational potential licensee would pay any royalties for a patent license if the patent owner did not have the ability to bring a lawsuit to seek a remedy. As a consequence, the legal structure for the determination of reasonable royalty damages in particular can play an important role in the royalties that are negotiated in patent license agreements. This is why patent owners with a strong focus on licensing their patent portfolios, whether SEPs or not, are concerned about recent developments in the patent damages law, including the Federal Circuit’s smallest salable patent practicing unit (SSPPU) principle,⁴ as well as the IEEE’s reference to the SSPPU in its recently revised IPR policy.⁵

The SSPPU principle can help ameliorate certain problems that arise in calculating damages in patent infringement litigation. First, the Federal Circuit explained that the SSPPU principle is an

¹ Bronwyn H. Hall, *Patents and Patent Policy*, 23 OXFORD REV. ECON. & POLICY 568, 572 (2007).

² See, e.g., *id.* at 574–75; Petra Moser, *Patents and Innovation: Evidence from Economic History*, J. ECON. PERSP., Winter 2013, at 23.

³ See, e.g., Dennis Carlton & Allan Shampine, *Identifying Benchmarks for Applying Non-Discrimination in FRAND*, CPI ANTITRUST CHRON. (Aug. 2014); Suzanne Michel, *Bargaining for RAND Royalties in the Shadow of Patent Remedies Law*, 77 ANTITRUST L.J. 889 (2011).

⁴ Also referred to as “smallest salable unit.” *Ericsson Inc. v. D-Link Sys., Inc.*, 773 F.3d 1201, 1227 (Fed. Cir. 2015).

⁵ Referred to as “smallest saleable Compliant Implementation.” IEEE-SA Standard Board Bylaws at 2 (2015), available at <http://standards.ieee.org/develop/policies/bylaws/approved-changes.pdf>.

evidentiary rule, adopted to avoid prejudicing a jury with large royalty base figures that result from the sales of the downstream product containing the SSPPU component.⁶ Second, starting with the SSPPU component also simplifies the apportionment problem in those situations where apportionment is used to calculate the reasonable royalty (as opposed to, for example, a “comparable license” approach). For example, apportioning the value of a Wi-Fi chip is substantially less complex than apportioning the value of a smartphone that contains the Wi-Fi chip.⁷ The Federal Circuit has explained that the reasonable royalty for a patent must appropriately apportion the value of the product between the patented technology at issue and the other contributions to the product’s value (including other patented technologies, non-patented technologies, other intellectual property, development activities, brand name, marketing, etc.).⁸

There are important nuances to the SSPPU requirement in the U.S. case law. Under the “entire market value rule” (EMVR) exception, the Federal Circuit allows the value of the downstream product to be used as the starting point for apportionment in situations where the asserted patent can be demonstrated to be the driver of demand for the downstream product.⁹ In addition, the Federal Circuit has stated that the SSPPU principle does not preclude the use of an existing license agreement as a “comparable” for determining the reasonable royalty if the agreement happens to use something other than the SSPPU as the royalty base.¹⁰

Despite these nuances, the SSPPU principle and the EMVR exception may not provide enough flexibility. Patents that are important drivers, but not the only driver, of downstream product demand (and thus do not qualify for the EMVR exception) may be undercompensated using an SSPPU apportionment approach that fails to attribute any of the downstream product value to these patents.¹¹ Given the skewed distribution of patent values, the percentage of patents that might fall in the category of “important, but not important enough to qualify for the EMVR exception” may be relatively small, but their share of patent value could be relatively large, making it important to properly align the relative incentives to produce such innovations. Rather than

⁶ *Ericsson*, 773 F.3d at 1226–27. See also Joseph Kattan, *The Next FRAND Battle: Why the Royalty Base Matters*, CPI ANTITRUST CHRON. (Mar. 2015). While it may be argued that, because the SSPPU is “only” an evidentiary rule in litigation it should not be used by parties in patent licensing negotiation, this ignores the point that a potential licensee will not pay more than the reasonable royalty determined in litigation. If the latter is based on the SSPPU principle, the potential licensee’s bargaining position rationally may be based on the SSPPU principle.

⁷ Another issue when using an apportionment approach is defining and measuring the value of the product that is to be apportioned. Under some conditions, the price or profit of the product may fail to fully reflect the value of the technology at issue if, for example, that value has been competed away. On the other hand, particularly in the context of standards, implementers likely anticipate having to pay royalties; in that case, the competitive price and profit will account for such expectations. An alternative starting point for apportionment, which is possible in some situations, is to identify an overall royalty burden. See, e.g., Gregory K. Leonard & Mario A. Lopez, *Determining Royalty Rates for Standard Essential Patents*, ANTITRUST, Fall 2014, at 86, 89–91.

⁸ *Ericsson*, 773 F.3d at 1226.

⁹ *LaserDynamics, Inc. v. Quanta Computer, Inc.*, 694 F.3d 51, 67–68 (Fed. Cir. 2012).

¹⁰ *Ericsson*, 773 F.3d at 1228.

¹¹ An important issue here involves which party in the supply chain the patent owner decided to sue. If the patent owner sues the SSPPU component supplier, it may be that this supplier is not able to capture downstream product value and thus the reasonable royalty owed by this supplier should not include such value. In general, the value of a product is not captured entirely by the seller of the product. Even in the case of a monopolist (with only a limited ability to price discriminate), a substantial amount of value typically is captured by entities further down the supply chain, including end users. Moreover, it is important not to confuse the value of an SEP with the value of a product feature that implements the SEP’s technology. For example, even where the product feature is valuable, the SEP itself may have little value if the feature could have been achieved just as readily using an alternative technology.

attempting to undo the SSPPU principle, which would risk swinging the pendulum back in favor of overcompensating minor patented innovations, the public policy and legal debate should focus on ensuring adequate compensation for this class of major patented innovations.

The apportionment problem may be even more difficult for an SEP than for a non-SEP that covers, say, a pharmaceutical product because of the large number of SEPs that are associated with standards and the incorporation of multiple standards in a given product. Moreover, given the large portion of product value that is generated by contributions other than the SEP at issue in such a context, the scope for hold-up may be greater in the SEP context than in a non-SEP context.¹² Finally, the potential for royalty-stacking concerns similarly may be greater in the SEP context than the non-SEP context. The SSPPU principle can address all of these concerns by limiting the amount of value that can be assigned to a given SEP. However, again, the SSPPU principle may be too inflexible to appropriately compensate certain high value SEPs.

The IEEE recently changed its IPR policy to recommend that the SSPPU principle be used to determine “reasonable and non-discriminatory” (RAND) royalties for SEPs. The IEEE policy appears to be somewhat more restrictive than the U.S. case law because it does not provide for an EMVR exception.¹³ On the other hand, the IEEE policy states that a reasonable royalty determination “should include, but need not be limited to, consideration of” the value contributed by the SEP’s technology to the SSPPU, which seems to imply some flexibility.¹⁴ Nevertheless, several SEP owners have reacted strongly to the change. For example, Qualcomm has announced that it will generally reconsider its relationship with the IEEE and that it will not make a licensing commitment under the new IEEE policy. Instead, it will offer alternative commitments on a case-by-case basis.¹⁵ From a policy-making perspective, the change to the IEEE IPR policy may constitute a “natural experiment” that provides useful information about the effects of policy changes on behavior related to licensing negotiations and SSO participation.

Portfolio Licensing

While patent licensing negotiations are influenced by litigation costs and potential litigation outcomes, there are important differences between a patent infringement lawsuit and a patent licensing negotiation. For example, in a typical patent infringement litigation, the patent owner is limited to asserting a relatively small set of patents, while in a patent licensing negotiation a large portfolio of patents may be in play.

Patent owners with large SEP portfolios argue that they are disadvantaged in patent licensing negotiations by the inability to litigate more than only a small set of patents at a time, so that in principle they have to bring multiple lawsuits to get “full value” for their portfolio. While this may be true, the patent system treats an SEP portfolio owner the same way as a patent owner with only one patent: before having a right to any compensation for use of any patented technology, a

¹² “Hold-up” in the SEP context occurs when an SEP owner exploits the fact that the standard is locked into the technology of the SEP to extract a royalty from licensees that exceeds the value of the SEP’s contribution. Lock-in, in turn, can arise because, for example, it is costly to switch out the technology of the SEP for another technology after the standard has been set and implementers have started selling products compliant with the standard.

¹³ It may be that it was considered unlikely that the EMVR exception would apply in the context of an IEEE standard. Nevertheless, the IEEE policy has even less flexibility than the U.S. case law, which as I have discussed could lead to under-compensation for high value patents.

¹⁴ IEEE-SA Standard Board Bylaws, *supra* note 5, at 2.

¹⁵ *Qualcomm Responds to Updated IEEE Standards-related Patent Policy*, EVALUATION ENGINEERING (Feb. 11, 2105), <http://www.evaluationengineering.com/2015/02/11/qualcomm-responds-to-updated-ieee-standards-related-patent-policy/>.

patent owner first must prove infringement and give the defendant the opportunity to prove patent invalidity.¹⁶ Moreover, potential licensees also would incur costs to litigate multiple cases and thus have similar incentives, all else equal, to negotiate a portfolio license so as to avoid such costs. Indeed, this often happens in practice.

Nevertheless, having an alternative dispute resolution mechanism for establishing the royalty for a portfolio of patents would seem to be valuable. This is particularly true in cases where the patent owner has a RAND commitment and therefore generally one would presume that a license is desired by both parties and the area of dispute relates to the terms of the license and, in particular, the royalty terms. At least one proposal has been made in this regard.¹⁷ However, such a mechanism would face substantial obstacles. Valuing a portfolio directly would require an examination of the underlying patents to determine their validity, essentiality, infringement, and contribution to product value. Perhaps this examination could be limited to a small subset of patents if such a subset dominates the portfolio's value. But, then, it is not clear what the advantages are of the alternative dispute resolution over traditional patent litigation.¹⁸ Other approaches to valuing a portfolio that could be used in an alternative dispute resolution framework include using comparable licenses as benchmarks and determining a total royalty burden for the standard and apportioning a piece of that burden to the portfolio in question.¹⁹

Valuation of SEPs

One element of the SEP debates is whether a "reasonable" royalty for an SEP under a RAND commitment should be based on a valuation of the SEP before the standard was set, often termed to be an "ex ante" valuation. Some have pointed out that such a valuation is ex ante only with respect to the implementer's sunk cost investments in developing their products; it is ex post with respect to the inventor's sunk cost investments in developing the patented invention.

However, the framework of the U.S. patent system calls for patents to be valued in the ex ante sense defined above. In the classic example of an inventor approaching a manufacturer with a patented invention that could improve the manufacturer's product, the negotiation between the inventor and manufacturer takes place after the inventor has developed its invention, but before the manufacturer has invested in adding the invention to its products. At the time the inventor decided to invest in innovation, the reward was uncertain, and driven by the expectations of the value that the innovation would generate. Similarly, in the pharmaceutical industry, a common pattern of drug development is for one firm to engage in research and development to identify a potential drug, which is then protected by a patent, and then to license a second (typically larg-

¹⁶ A further complication is whether a patent that is declared by its owner to be an SEP is actually "essential" to practice the standard in question.

¹⁷ Mark A. Lemley & Carl Shapiro, *A Simple Approach to Setting Reasonable Royalties for Standard-Essential Patents*, 28 BERKELEY TECH. L.J. 1135 (2013).

¹⁸ In my experience, when deciding which patents to assert, patent owners with portfolios put the most emphasis on identifying those patents for which they have the greatest probability of establishing liability, with the expected damages (conditional on liability) a secondary consideration.

¹⁹ A number of possible approaches to apportioning an overall royalty burden among portfolios exist. For example, an apportionment could be based on an assessment of the relative technical contributions of the portfolios. This approach was taken by the court in *Innovatio*. See *In re Innovatio*, 921 F. Supp. 2d 903, 1072 (2013). Another possibility is to analyze indicators of patent value, such as the number of forward citations a patent receives. Forward citations have been demonstrated in the economics literature to be predictor of patent value. See, e.g., Manuel Trajtenberg, *A Penny for Your Quotes: Patent Citations and the Value of Innovations*, 21 RAND J. ECON. 172 (1990).

er) firm to further develop and market the drug. The license is negotiated after the first firm has incurred its development expenses, but prior to the second firm having incurred its expenses.

Such an ex ante framework is also consistent with how non-SEPs are treated under U.S. case law. A reasonable royalty in patent infringement litigation typically is determined with reference to a “hypothetical negotiation” set just prior to infringement, which is after the inventor’s investments and before at least some of the manufacturer’s investments. Moreover, the Federal Circuit has made clear that the reasonable royalty should be limited to the contribution of the invention to the product’s value, after properly apportioning out other contributions to the value of the product. Georgia-Pacific Factor 13 specifically identifies contributions of the implementer,²⁰ which in many instances will have been made prior to the hypothetical negotiation. In contrast, there appears to be no basis in the law for the reasonable royalty to account for the investments incurred by the inventor in developing the invention. If SEPs are to be treated like other patents, the ex ante framework should apply.

If SEPs are to be treated like other patents, the ex ante framework should apply.

Any presumption that an SEP must be valuable because the technology it covers was chosen for inclusion in the standard should be avoided. First, as the Federal Circuit has specified, it is important to distinguish the value of the patented technology itself from the value created by standardization that could have been achieved using an alternative technology.²¹ Second, there is some evidence that, prior to being included in a standard, SEPs tend to have lower indicators of value than non-SEPs.²² Third, standardization is needed in situations where there are alternative technologies that could be used, but for the purposes of achieving interoperability, one must be chosen. The existence of alternatives reduces the value of any one of them (relative to the absence of such alternatives). Thus, a given SEP, despite its technology having been chosen, may actually be of low (ex ante) value.²³

An illustrative example of the importance of alternatives in ex ante valuation occurred during the development of the 802.11g amendment to the Wi-Fi standard. A company called Intersil submitted a proposal to the 802.11g working group that included a particular technology called CCK-OFDM, while Texas Instruments submitted a proposal that included an alternative technology called PBCC.²⁴ There was substantial support for both proposals in the working group.²⁵ Texas Instruments offered to license its intellectual property related to PBCC on a royalty-free basis.²⁶

²⁰ Georgia-Pacific Corp. v. U.S. Plywood Corp., 18 F. Supp. 1116, 1120 (S.D.N.Y. 1970).

²¹ Ericsson, 773 F.3d at 1232.

²² See, e.g., Rudi Bekkers et al., *Selected Quantitative Studies of Patents in Standards* 30–32 (Hitotsubashi Univ. (Tokyo), Inst. of Innovation Research, PIE/CIS Working Paper 626, June 2014), available at <http://cis.ier.hit-u.ac.jp/Common/pdf/dp/2014/dp626.pdf>.

²³ For illustrative purposes, suppose that the population of patents consists of a large number of pairs of alternative patents that are “good,” but not perfect, substitutes for each other plus a small group of patents that each have no good alternative. The distribution of patent value for this population would be highly skewed with a large number of patents (those with an alternative) having relatively low value and a small number of patents (those without an alternative) being of high value. If the SSO made technology choices based on value alone and restricted itself to setting a standard only with regard to technologies for which there were alternatives, the resulting SEPs (i.e., patents with alternatives that were chosen for inclusion) would have greater average value than the patents with alternatives not chosen, but less average value than patents overall.

²⁴ Steven J. Vaughan-Nichols, *What Is PBCC Anyway?* WI-FI PLANET.COM (Oct. 8, 2002), <http://www.wi-fiplanet.com/columns/article.php/1478441/What-is-PBCC-Anyway.htm>.

²⁵ *Id.*

²⁶ <https://mentor.ieee.org/802.11/dcn/01/11-01-0023-01-0000-texas-instruments-ip-statement-for-802-11-tgg.ppt>.

Intersil then similarly offered to license its SEPs on a royalty-free basis if its proposal was adopted for 802.11g.²⁷ Ultimately, the Intersil proposal was adopted, but only after political compromise.²⁸ This episode demonstrates that the existence of alternatives can create competition that drives down RAND royalty rates (in this case to zero).²⁹ On the other hand, in the absence of alternatives, an SEP may command a relatively high RAND royalty rate. Thus, it would also be incorrect to presume that an SEP is of minimal value, or should be licensed royalty-free.

The risk of overcompensation (due, e.g., to hold-up) is a greater concern with regard to SEPs with relatively low value, while the risk of under-compensation (due, e.g., to consequences of legal rules like the SSPPU principle) is a greater concern with regard to SEPs with relatively high value. The former are likely to be substantially greater in number, but the latter are more likely to contribute substantially to social welfare. It is difficult to create simple rules that fully address concerns about ensuring an overall royalty burden for downstream products that appropriately reflects the value of the technologies used while still providing appropriate incentives to inventors and manufacturers to innovate.

Is Hold-Up a Real Problem?

Another point of contention in the SEP debates is whether hold-up by SEP owners is a real problem, or just a theoretical one. Part of the debate is determining what constitutes a “real” problem. In particular, is hold-up a real problem only if there is evidence of standards being derailed by SEP owners or widespread instances of exorbitant royalties being charged? Such criteria would be too stringent. We would not expect to see SEP owners derailing standards because that would decrease their royalty streams. The situation is analogous to that of a traditional monopolist; the monopolist would set its price so as to exercise market power, but not so high as to entirely eliminate demand. Thus, the existence of a functioning market does not mean there is no exercise of market power. Nor would we expect to see widespread instances of exorbitant royalties being charged because SSOs have adopted RAND commitments, disclosure requirements, and other intellectual property policies precisely to prevent such outcomes, and many SSO participants in fact abide by these policies. Analogously, crime is not particularly widespread, and yet it is a real concern as reflected in laws prohibiting crime. For those SSO participants that do not abide by the policies, it likely is in their best interest to charge royalties that are not obviously exorbitant, even if above the RAND level.

In any event, there is evidence that is suggestive of hold-up or attempted hold-up. Companies have been found to have violated their RAND commitments in legal proceedings.³⁰ Companies have sought royalties that exceeded the RAND level as determined in subsequent legal pro-

²⁷ http://standards.ieee.org/about/sasb/patcom/loa-802_11g-intersil-01Jul2002.pdf; http://standards.ieee.org/about/sasb/patcom/loa-802_11g-intersil-04Apr2003.pdf.

²⁸ See Vaughan-Nichols, *supra* note 24.

²⁹ In a recent paper, Josh Lerner and Nobel Prize winner Jean Tirole demonstrate that efficient outcomes can be obtained if patent owners make “structured price commitments,” i.e., if they noncooperatively announce their royalty rates, prior to the standard being set. Josh Lerner & Jean Tirole, *Standard Essential Patents*, J. POL. ECON. (forthcoming 2015), available at http://www.tse-fr.eu/sites/default/files/medias/doc/wp/fit/wp_tse_441_v3.pdf. Intersil’s and Texas Instruments’ licensing announcements are similar to Lerner and Tirole’s structured price commitments.

³⁰ See, e.g., Order Granting Partial MSJ and Denying Motion to Stay at 15, *Realtek Semiconductor Corp. v. LSI Corp.*, No. C-12-3451 (N.D. Cal. May 20, 2013).

ceedings by an amount that, together with the particular negotiating tactics they employed, suggest an attempt at hold-up.³¹

It is, however, often difficult to distinguish between a patent owner attempting “deliberate” hold-up (i.e., engaging in bad faith efforts to obtain a royalty above the RAND level) and a patent owner using available bargaining leverage to obtain what it in good faith believes to be the RAND royalty. Thus, a royalty above the RAND level can result even in the absence of bad faith hold-up, and the use of various forms of bargaining leverage, such as the threat of an injunction, is not necessarily inconsistent with a RAND commitment. As with other issues discussed here, there are no simple answers.

Injunctions, Bargaining Leverage, and Hold-Out

The threat of an injunction is the form of bargaining leverage that has perhaps been most associated with hold-up. An injunction can, indeed, provide powerful leverage to the patent owner, particularly when the patented feature is only a small part of the potential licensee’s product, so that the loss to the potential licensee from an injunction far exceeds the RAND royalty for the patent.

Where things stand currently on injunctions in the U.S. federal courts is probably about right: the existence of a RAND commitment should enter into the eBay four factor test³² and weigh against an SEP owner getting an injunction (because the SEP owner has acknowledged that it is willing to license the patent), but at the same time the SEP owner should not be entirely barred from getting an injunction, for example, in the case of a potential licensee unwilling to pay a royalty that has previously been determined (perhaps by a court) to be RAND. Nevertheless, even in the situation of an “unwilling” licensee, the SEP owner that has been granted an injunction should be constrained to extracting only the RAND royalty (plus any other remedy available under applicable law).³³

A concern for SEP owners is that the greater difficulty in obtaining an injunction, and the concomitant reduction in bargaining leverage, creates a “costless option” for a potential licensee: negotiate in bad faith in an attempt to obtain a royalty below the RAND level, with the worst case outcome being the RAND royalty if litigation results. This has been termed “hold-out.” Bad faith on the part of the potential licensee should also be considered to be inconsistent with RAND.³⁴ However, an SEP owner still has other forms of bargaining leverage available.³⁵ For example, the

³¹ For example, Innovatio sued small businesses that offered Wi-Fi to customers, such as individual hotels, seeking royalties of over \$2000 per location and threatening expensive litigation. See <http://patentexaminer.org/2011/09/innovatios-infringement-suit-rampage-expands-to-corporate-hotels/>. The court-determined RAND royalty for the Innovatio portfolio was approximately \$0.10 per Wi-Fi chip. Innovatio, 921 F. Supp. 2d at 1090.

³² eBay Inc. v. MercExchange, LLC, 547 U.S. 388, 391–94 (2006).

³³ The rules governing the issuing of injunctions differ in other venues. For example, in a U.S. International Trade Commission proceeding, where an “exclusion order” is the primary available remedy, the Commission must issue an exclusion order unless “after considering the effect of such exclusion upon the public health and welfare, competitive conditions in the United States economy, the production of like or directly competitive articles in the United States, and United States consumers, it finds that such articles should not be excluded from entry.” See 19 U.S.C. § 1337(d)(1). Some of the public interest factors that the ITC must consider appear to an economist to overlap with the eBay factors, but the Federal Circuit has stated that the two tests are not identical. See Spansion, Inc. v. ITC, 629 F.3d 1331, 1359 (Fed. Cir. 2010).

³⁴ This statement is my view as an economist and is not meant to reflect any legal principle.

³⁵ These forms of leverage may still allow an SEP owner to achieve hold-up. For example, Melamed discusses how a patent owner may extract above RAND royalties from small potential licensees using litigation costs as leverage, and then use these royalties as “benchmarks” in an attempt to obtain above RAND royalties in subsequent licensing negotiations. See A. Douglas Melamed, Prepared Statement for the Record for Intel Corporation, Before the Senate Comm. on the Judiciary Subcomm. on Antitrust, Competition Policy and Consumer Rights, on Standard Essential Patent Disputes and Antitrust Law (July 30, 2013), available at <http://www.intel.com/content/dam/www/public/us/en/documents/corporate-information/melamed-testimony-july-30-2013-statement.pdf>.

costs of litigation may weigh more heavily on the potential licensee than the SEP owner (for example, the existence of litigation may cause some of the potential licensee's customers to change suppliers), "mistakes" in litigation damages awards are likely to be greater on the upside than the downside,³⁶ and the royalty after a finding of validity and infringement can exceed a royalty negotiated prior to such a finding. Moreover, the potential licensee may have limited bargaining leverage in many cases. For example, absent capital market imperfections, the potential licensee cannot create leverage through delaying payment of the RAND royalty if the patent owner can collect appropriately measured prejudgment interest. Similarly, in most cases the potential licensee's litigation costs are not lower than the SEP owner's and likely are greater. The one area where the potential licensee could have bargaining leverage is the risk to the SEP owner's future licensing endeavors if its litigated patents were found to be invalid or not infringed in a lawsuit against the potential licensee.³⁷ However, such bargaining leverage may be mitigated if the SEP owner has a large portfolio.

In general, it is difficult to distinguish between hold-up or hold-out, on the one hand, and hard negotiating, on the other.

In general, it is difficult to distinguish between hold-up or hold-out, on the one hand, and hard negotiating, on the other. The determination of the RAND royalty is a complex task that is subject to substantial disagreement between the parties. In that context, both parties are entitled to vigorously argue their positions in good faith. However, even with good faith on both sides, agreement may not be reached. One or both parties may then seek a litigated resolution. Failure to reach an agreement or the bringing of litigation do not, in isolation, indicate bad faith or a violation of a RAND commitment. Something more is needed. An example is if the SEP owner sues and seeks an injunction against an alleged infringer without ever making a licensing offer. It is difficult to see how such behavior is consistent with a RAND commitment.

Conclusion

The various debates surrounding SEPs are too important to be left to rhetoric and sound bites without solid economic basis because any policies that emerge from these debates could have substantial effects on standardization and innovation, which have been important contributors to social welfare. There are no simple answers to many of the questions that have been raised because of the complexities of the underlying economic systems. In that context, the development of sound policy requires that the complexities of the issues and resulting ambiguities are understood and properly considered, even when they cannot be fully taken into account. ●

³⁶ Damages are generally bounded below by zero, which limits the size of downside mistakes. While damages can effectively be negative if, for example, a judge awards the defendant its costs, this outcome is rare in the U.S. court system. At the same time, damages are unbounded above. Thus, in principle, the size of upside mistakes are not limited, although the larger the mistake the more likely it is that it would at least be partially corrected via motion practice or on appeal.

³⁷ Note that it would be consistent with RAND for a potential licensee to argue that a negotiated royalty should be lower to reflect the uncertainty that the patent is valid and infringed. The leverage referred to here involves the potential licensee seeking a lower royalty on the basis that, with a settlement, the SEP owner does face any risk to its licensing revenue stream from other licensees.

Paper Trail: Working Papers and Recent Scholarship

Editor's Note: Editor Bill Page reviews a paper by Jeremy Bertomeu, John Harry Evans III, Mei Feng, and Ayung Tseng addressing the anticompetitive effect of information sharing among firms, using the automobile industry as test of the generated hypotheses. Editor John Woodbury considers a paper by Gregory J. Werden challenging the outcomes of retrospective studies of mergers and another paper by Justin P. Johnson offering a benign explanation for “loss leaders” when consumers are characterized by bounded rationality.

Send suggestions for papers to review to. or

—WILLIAM H. PAGE AND JOHN R. WOODBURY

Recent Papers

Jeremy Bertomeu, John Harry Evans III, Mei Feng & Ayung Tseng,
Tacit Collusion and Voluntary Disclosure: Theory and Evidence from the U.S. Automotive Industry,
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2613970

This paper studies how rivals can share information to facilitate tacit collusion.¹ The authors offer both theory and empirical evidence in their analysis. The theoretical model is of two firms that compete over many time periods by, first, deciding whether to share information (“the information decision”) and then choosing outputs (“the operation decision”). They see first a “public signal,” such as a forecast of overall economic activity, then a “private signal” of their own forecast of demand in the market. At that point, knowing that both of those factors will affect the eventual market price, they decide whether to share the private signal with their rival—providing the information to some intermediary like a trade association, on the condition that it will distribute it only if both rivals opt to share. Only after the rivals receive the shared information (or not) do they make their profit-maximizing output decisions for the period.

If the rivals only competed in one time period, no firm would ever share its private information, because its rival would use it to estimate the firm’s output choice then opportunistically choose a higher output, thus lowering the firm’s expected profit. If the firms compete over many periods, however, the firms would recognize that choosing the higher output would damage their reputation and likely provoke retaliatory price cuts (the “punishment path” or the “Nash equilibrium of the

¹ The authors use “tacit collusion” in the usual economic and legal sense of interdependent noncompetitive pricing. *See, e.g.*, *Brooke Group Ltd. v. Brown & Williamson Tobacco Corp.*, 509 U.S. 209, 227 (1993) (Tacit collusion is “the process, not in itself unlawful, by which firms in a concentrated market might in effect share monopoly power, setting their prices at a profit-maximizing, supracompetitive level by recognizing their shared economic interests and their interdependence with respect to price and output decisions.”). The authors use the term “tacit agreement” to describe an instance of tacit collusion involving the information sharing they describe. Although the Supreme Court continues to include tacit agreement (but, confusingly, not tacit collusion) within the scope of Section 1 of the Sherman Act, *Bell Atlantic Corp. v. Twombly*, 550 U.S. 544, 553 (2007) (“[T]he crucial question is whether the challenged anticompetitive conduct stem[s] from independent decision or from an agreement, tacit or express.”), the authors of the paper under review do not use the term in any legal sense. They recognize, however, that the mechanisms of information sharing and their effectiveness in maintaining a noncompetitive price level are relevant to the Sherman Act analysis.

single-period game”) in later time periods. Consequently, both firms would have a greater incentive to share their private information and use it to increase joint profits by tacit collusion. The authors recognize that this process could involve cheating, punishment, and then renegotiation, but such a scenario might require direct communications that would attract the interest of antitrust enforcers.

A key implication of the model is that, at lower market demand levels, firms can tacitly coordinate on the profit-maximizing output by sharing the forecast information. At high demand levels, however, the incentive to cheat on a tacit arrangement to set jointly profit-maximizing outputs would be greater because the short-term returns from opportunistically setting a higher output would be greater; consequently, the incentive to cheat would overwhelm the expected cost of future punishments, and the agreement would break down. At higher demand levels, the firms would have to settle tacitly on a somewhat lower, but “incentive-compatible” combined output.

If the firms do not share information, then they can only act on the public macroeconomic forecast and their own private demand forecasts, so they will lack the information necessary to achieve the full profit-maximizing output by tacit collusion. When their estimated demand is low, based on the public signals and their private signals, the firms will always share information and maximize joint profits. But if the demand is higher than some threshold level, they will anticipate that the incentive to cheat will be correspondingly higher, so they will not share information—the expected profit from doing so will actually be lower because of the incentive to increase output opportunistically using the shared information. Thus, they will settle for the profit in the single-period game without information sharing.

The authors’ model implies that the degree of information sharing should depend in substantial part on industry demand and on the extent of the rivals’ focus on short-run profit. Using historical data from the American auto industry, they authors conducted regressions to test these hypotheses. The industry was well-suited to the study, at least until the mid-1990s, because it was fairly concentrated among the big three domestic producers—between 60–90%. Moreover, the means of information sharing was similar to those hypothesized by the theory. The automakers observed the level of macroeconomic activity each month, made their own forecasts, and then chose whether to submit their own forecasts to Ward’s Automotive Report, which would then publish the information in weekly newsletters to their subscribers, which included both automakers and dealers, only if all three submitted forecasts. The forecasts they submitted could also be relatively more accurate, if a producer invested more in information-gathering, then submitted revised forecasts during a given month to be distributed in a later week’s newsletter. The automakers’ actual production each month then revealed the accuracy of their predictions.

The authors use frequency, time horizon, and accuracy of production forecasts to measure the degree of information-sharing. Their proxies for the firms’ degree of near-term focus are the rate of turnover of the CEO of the producer (a high rate would tend to make the producers more focused on short-term profit) and the Altman z-scores² for the firms (a measure of financial distress, which would also tend to focus the firm on the near-term bottom line). Their proxy for automotive demand is the monthly unemployment rate, which is apparently a very good predictor of new-car sales. They also take account of a variety of control variables that could also affect the automakers’ incentives to share information.³

² Edward I. Altman, *Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy*, 23 J. FIN. 589 (1968).

³ These included capacity utilization, inventory level, production volatility, inflation, and the prices of inputs like steel and electric power.

The findings of the study generally confirmed both of the paper's hypotheses of the determinants of information-sharing. The authors found significant correlations between their measures of the degree of information sharing and their measures of demand (the unemployment rate) and short-term focus (CEO turnover and financial distress). They conclude that their findings were "generally consistent with automobile manufacturers communicating with each other more intensely, reflected in more frequent production forecasts and longer forecast horizons, when the expected demand in the industry is lower and when the firms face longer decision horizons." (p. 21.) In a later section, they also show that, when unemployment is high, the producers tend to revise their production forecasts and their actual production over the course of the month to take account of their rivals' forecasts distributed during the same month.

The paper also includes an extensive and illuminating review of both the empirical and theoretical literature on disclosure and information sharing under different competitive conditions.

The authors suggest that "the intuition" of their results "would appear to generalize to other industries in which there is evidence of tacit collusion," (p. 5) citing both other studies and three antitrust cases.⁴ They also suggest that their results "have potential regulatory implications," because they "show how information sharing under tacit agreement can potentially facilitate coordination of total industry production to the detriment of consumers," (p. 5.). They cite two information-exchange cases from the 1920s as "document[ing] early antitrust actions against information sharing in trade associations."⁵

I would add that courts view exchanges of predictions of *future* competitive conduct, like the automakers' production forecasts this paper studies, more suspiciously than, for example, exchanges of *present* prices, but the (relatively) public distribution of the forecasts in a trade journal would almost certainly have insulated the automakers from liability.⁶

—WHP

Gregory J. Werden, Inconvenient Truths on Merger Retrospective Studies (Jan. 5, 2015), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2545343

In a recent issue of *The Antitrust Source*, Robert Skitol reviewed a book by John Kwoka assessing retrospective merger analyses.⁷ Skitol notes that "Kwoka undercuts any confidence that the agencies have effectively protected consumers from the anticompetitive effects of merger activity." (Review p. 1.) Skitol quotes Kwoka as concluding that the net effect of enforcement policy "has been to focus on mergers most directly causing harm, but the diminished attention to mergers involving somewhat lower market shares and concentration appears to have resulted in approval of significantly more mergers that prove to be anticompetitive."⁸ As Skitol notes, one of Kwoka's

⁴ The citations are incomplete, but they apparently refer to *United States v. Container Corp.*, 393 U.S. 333 (1969); *United States v. Apple Inc.* 952 F. Supp. 2d 638 (S.D.N.Y. 2013), *aff'd*, Nos. 13-3741-cv, 2015 WL 3953243 (2d Cir. June 30, 2015); *United States v. AU Optronics Corp.*, 09-CR-0110 SI, 2012 WL 2120452 (N.D. Cal. June 11, 2012).

⁵ *Id.* They cite *Maple Flooring Mfrs. Ass'n v. United States*, 268 U.S. 563, 584 (1925) and *Am. Column & Lumber Co. v. United States*, 257 U.S. 377 (1921).

⁶ See William H. Page, *Signaling and Agreement in Antitrust Law* (June 19, 2015), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2620786. An earlier version of this paper is forthcoming in *CONCURRENCES J.* (2015).

⁷ Robert Skitol, *Book Review, A Harsh Report Card on the Merger Enforcement Process*, ANTITRUST SOURCE (Feb. 2015), http://www.americanbar.org/content/dam/aba/publishing/antitrust_source/feb15_full_source.authcheckdam.pdf.

⁸ *Id.* at 2 (citing JOHN KWOKA, *MERGERS, MERGER CONTROL, AND REMEDIES: A RETROSPECTIVE ANALYSIS OF U.S. POLICY 155* (2015) [hereinafter KWOKA]).

conclusions is that based on retrospectives, the “agencies should develop ‘a better understanding of the a priori characteristics of incorrectly cleared mergers, and hence the sources of policy errors.’”⁹

In a recent paper entitled “Inconvenient Truths on Merger Retrospective Studies,” Gregory J. Werden, Senior Economic Counsel in the Justice Department’s Antitrust Division, effectively tells us “not so fast.” As Werden notes (p. 2), most of the retrospective merger studies rely on the “difference-in-difference” (DiD) methodology. These are studies that focus on whether a consummated merger has resulted in market power as evidenced by post-merger price increases. More specifically, the DiD methodology compares the changes in prices for products/services offered by the merged firm (the “treatment” group) as compared to changes in the prices of a control group of products/services. If the merger led to market-power driven higher prices, then the change in price for the products/services affected by the merger will exceed those in the control group.

Werden makes two overarching criticisms of the DiD merger retrospective approach. First, he explains why the DiD retrospectives are a weak foundation for any evaluation of the price effects of mergers. Second, he argues that, in any event, the retrospectives cannot inform sensible changes in merger enforcement policy. But he also suggests the possibility that these analyses can be more informative if conducted in the context of a case study (although even here, Werden seems skeptical).

“Inconvenient Truths” in the Use of DiD Analyses

As most researchers are well aware, the single most significant choice in conducting a DiD analysis is identifying a suitable control group. If the choice of the control group is incorrect, the results of the DiD analysis are worthless. That is, comparing the changes in the prices of the merged firm to those in a flawed control group would be a meaningless exercise. If the price changes for the merged firm were greater than those in the incorrectly-defined control group, one could not reasonably conclude that the merger resulted in higher prices.

For example, in a broad product market, using the prices of products that are similar to but not produced by the merging parties as the control is not likely a reasonable choice. If the merger affects prices of the merging firm’s output, it’s likely to affect the prices of other “like” products produced by other firms. Werden seems to suggest that if markets are “local,” products sold in locations not affected by the merger could serve as a more suitable control, although the different locales must still face identical demand and cost conditions. (note 14.) While frequently DiD analyses will include supply and demand variables as a way to account for differences in the treatment and control groups, it still must be true that “the unaffected prices must be influenced by exactly the same supply and demand forces that influence the affected prices” (p. 4)—something difficult to verify.

“As with most econometric work” (p.6), the researcher is confronted with the need to choose the data, the time period, the price measures and the statistical methodology for the investigation. Werden notes that “[f]ew results are so robust that no choice matters, and different researchers are apt to make choices different enough to produce substantially different estimates of merger effects.” (p. 6.) Werden provides a number of examples that illustrate this sensitivity. In one review of the Northwest/Delta merger, the researchers used a variety of control groups and found that the

⁹ *Id.* at 3 (citing *Kwoka*, *supra* note 8, at 159).

post-merger price increase ranged from 0 to 6 percent. Using an algorithm that identifies the best match between the control and treatment groups, the estimated fare increase was only 1 percent.¹⁰

The merger retrospective that focuses on price increases may be misleading if there are significant post-merger quality improvements that are available to consumers at a higher price. The retrospective will then conflate anticompetitive price increases with price increases resulting from improved quality. “While quality changes might be accounted for, that is difficult and unlikely to be attempted.” (p. 7.)

Werden notes that the magnitude of the price effects may depend on the time period chosen. While merger retrospectives typically focus on a one- or two- year period after the merger, he notes that “a few merger retrospectives have found long-term merger effects that differed materially from short-term effects.” (p. 7.) But then he goes on to argue that one really cannot assess the long-term effects of the merger: “It is not credible that a change in pricing five years after a merger is from the merger.” (p. 8.)¹¹

In addition, he observes that many of the retrospectives generate puzzling results, at best. For example, a retrospective of the Maytag-Whirlpool merger found significant price increases for dryers, but not for washers, even though the market shares for each product were about the same. “[I]f the study were taken at face value, an unknown force must have caused the merger to have dramatically different effects in nearly identical markets.” (p. 14.)

Werden concludes that “[a]n inconvenient truth is that econometric merger retrospectives cannot come close to definitively determining actual merger effects.” (p. 9.)

Inability of Merger Retrospectives to Inform Merger Enforcement Policy

Even if the DiD merger retrospectives were not characterized by the infirmities he describes, Werden concludes that “inconvenient truths preclude nearly every way of using econometric merger retrospectives to recalibrate merger enforcement.” (p. 9.) For example, Werden notes that using the retrospectives to adjust the HHI thresholds in the 2010 Merger Guidelines would not be particularly helpful “for the simple reason that market shares are determinative of neither agency assessments nor actual merger effects.” (p. 9.) And the retrospectives, while addressing price effects, do not address the agency “failures” in identifying the relevant antitrust market.

Werden also argues that any “meta” analyses of the retrospectives—relating the estimated price effects from the retrospectives to various factors—is a futile and potentially misleading research path: “Merger assessment is so heavily fact dependent that every case—or at least every close case—is unique, so the available data cannot trace out a general rule.” (p. 10.)

More generally, Werden takes retrospective advocates to task because they “have not explained what guidance they anticipate or how retrospectives provide it. Presumably, they intend something simple; for example, if merger retrospectives often estimate adverse merger effects, the guidance would be to tighten enforcement. But such guidance is too vague to be of much help.” (p. 12.)

¹⁰ Werden, p. 7 (citing Aditi Mehta & Nathan Miller, *Choosing the Appropriate Control Group in Merger Evaluations*, in THE PROS AND CONS OF MERGER CONTROL 189 (Swedish Competition Auth. 2012)), available at <http://www.microeconomix.eu/ressources/pros-and-cons-merger-control-0>. Werden does not opine on whether the use of this algorithm substantially mitigates the danger of picking the wrong control group.

¹¹ It is unclear, then, why Werden gives any weight to the long-term differences.

Increasing the Usefulness of Retrospective Analyses

While Werden apparently does not consider the retrospectives by themselves particularly useful, he does suggest how their antitrust utility can be increased. In addition to assessing price effects, retrospective studies “could ask whether postmerger competition conformed to the agency’s prior predictions, and if not, what specific aspects of the agency’s assessment were inaccurate. Identifying a source of systematic error in agency merger assessment could lead to a significant recalibration of enforcement” (p. 15.) In other words, a retrospective that looks more like a case study could prove useful. It would require that the researcher evaluate the reasoning behind the agency’s decisions, identifying the apparent source of the erroneous prediction as “some combination of faulty facts, erroneous economics, bad breaks, and random variation.” (p. 15.) Retrospectives that found a pattern of faulty facts or erroneous economics could result in significant modifications to the merger review process. While this approach requires detailed information on the merger (or other) assessment, Werden seems to agree with Dennis Carlton that the “record-keeping” of the agencies must become far more detailed and careful for the case-study approach to be feasible. (note 57.)

Closing Observations

I certainly agree that retrospectives combined with detailed information on the agencies decisions to (e.g.) clear a merger or not could prove very informative, although the data requirements will be substantial.

However, I believe Werden’s dismissal of retrospective studies themselves goes too far. Werden notes that “ *s with most econometric work*, merger retrospectives present myriad choices: data sources, price measures, time periods, and statistical methods. Few results are so robust that no choice matters, and different researchers are apt to make choices different enough to produce substantially different estimates of merger effects.” (p. 6) (emphasis added) Thus, in broad strokes, the issues raised by Werden with respect to retrospective analyses are the same as those raised for any econometric work, whether estimating price fixing damages or estimating a demand curve. This is not to counsel despair when using econometrics in antitrust (or any other field). Instead, we economists (and lawyers using economists) need to be reminded of the potential fragility of conclusions reached on the basis of econometrics and the need to test the robustness of the conclusions regardless of the specific analysis.¹²

Moreover, one can make lemonade out of the study “lemons.” Multiple DiD retrospectives on the same merger that yield different (or puzzling) results regarding the post-merger price changes is not in and of itself a bad thing. It provides researchers a platform to understand the sources of those differences and evaluating which choice combination is likely to yield a more “true” result. Indeed, this “dialectic” is part of the research process itself.

Certainly, Werden’s reminder is a cautionary note when deciding how much weight should be ascribed to the conclusions of retrospectives and his paper provides a useful guide to evaluate those studies.

¹² In my personal experience and in observing colleagues in the course of a merger matter or other litigation matter, the kiss of econometric death is to ignore robustness issues. That is, will relatively small changes in price measures (or other variables of interest), time periods, and statistical methods result in different econometric results? (It is unusual for any given statistical question to have multiple data sets available to the researcher.)

Justin P. Johnson, *Unplanned Purchases and Retail Competition* (May 5, 2014),
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2319929

While not as current as many of the papers typically considered in *the Per r il*, a paper by Justin P. Johnson, Associate Professor at the Samuel Curtis Johnson Graduate School of Management, Cornell University—“Unplanned Purchases and Retail Competition”—has apparently been making the rounds at the agencies and elsewhere. There are three interesting “hooks” in this paper: it assesses the practice of “loss leader” pricing; it accounts for bounded rationality of consumers (thus placing the paper broadly in the behavioral economics literature); and it finds that bounded rationality drives virtually all of the paper’s key results.

“Loss-leader” pricing is the practice by multi-product firms of selling “certain products beneath cost, with the hope of attracting additional customers who will also buy other, higher-margin items.” (p. 1.) Johnson notes that this practice has been banned in a number of countries and in about half of all U.S. states, explaining that “a serious concern of antitrust authorities is that larger firms often price below cost on the core product lines of smaller rivals.” (pp. 1–2.) The conclusion of the paper is that such practices need not be predatory in the presence of bounded rationality, and, indeed a ban on such practices could harm consumers.

More specifically, consumers are assumed to make “unplanned” purchases when they visit retail outlets, purchases that ex ante they believed they would make with a small probability: “It must be [for example] that consumers are (correctly) confident that they need some particular product, such as milk or bread (which will be priced below cost), but underestimate their tendency to purchase other goods.” (p. 2.) In Johnson’s model, a key parameter is what he calls the “accuracy ratio” that reflects the consumer’s biased belief on likely purchases. This is what the consumer ex ante believes his purchase probability of any good x is compared to the unbiased (actual) probability of purchase. If the ratio equals 1, the consumer has an unbiased and accurate assessment of the likelihood of the purchase of good x . A ratio less than one indicates that the consumer is underestimating the likelihood that she will purchase x . With respect to grocery stores, “staples” (like bread and milk) will have a high accuracy ratio compared to the accuracy ratio of unplanned purchases.¹³

In the absence of consumer bias (and when all rival firms carry the same array of products), Johnson shows that loss-leading does not occur. Since a purpose of loss-leaders (in Johnson’s model) is to encourage unplanned purchases on the higher-margin items, there is no payoff to the firms when there are no unplanned purchases: “Neither differing shapes of demand nor differing marginal costs by themselves lead to loss leading. Likewise, the fact that demand is stochastic and that there is rivalry among firms is not enough. Rather, if loss leading is to exist . . . it must be that consumers make unplanned purchases.” (p. 8.)

Further, in choosing which products are to be loss leaders, all that matters is the set of accuracy ratios. “A price cut is more effective at generating in-store traffic when customers expect that they will buy [the product],” noting that among grocers (for example) loss leaders tend to be staples, products that consumers purchase regularly. (p. 9.)

The results above hold when all firms are offering the same set of products. Johnson then considers a strategy of loss leaders when there are smaller firms that carry only a subset of the products carried by the larger firms. A possible anticompetitive concern in this case might be that the

¹³ Because the accuracy ratio will tend to be highest for goods or services that the consumers “need,” the unplanned purchases can be regarded as “discretionary” purchases.

larger firms will price below cost on those goods that are offered by the smaller firms so as to force their exit or otherwise impair their ability to serve consumers.¹⁴

The small firm carries as many staple goods as possible, and these are priced lower than non-staple goods. While the small firm could instead carry more non-staples sold at a higher margin, these non-staples are less effective in attracting customers because the consumer underestimates her purchases of these non-staples when choosing what outlet to patronize. While, in principle, the small firm could charge a higher price for the staples, the lower price of staples “increases the number of customers who shop [at the smaller firm].” (p. 16.) This is because consumer utility from shopping at the smaller firm is higher with the lower staples price than with the higher price.

Note that the consumer chooses which store—large or small—to patronize based on a comparison of ex ante expected utility. With consumers underestimating the “true” likelihood that they will purchase non-staples, consumers underestimate the “true” utility of shopping at the larger store. Even though the large firm may be pricing staples below cost, smaller firms can still survive charging a higher price for the staples (to at least break even) because such outlets may be more convenient (lower transport costs) for consumers than the larger firm.

Further, Johnson shows that the below-cost price being charged by the larger store is not a result of the subset of staples carried by the smaller store: “[W]hich goods are priced below cost by [the larger store] is driven entirely by the nature of consumer bias, not as such by the product line of its [smaller rival]. Similarly, [the smaller store] willingly chooses to carry goods that it knows the [larger store] is pricing below cost, where its decision is also driven by the extent of consumer bias across the different products.” (p. 17.)

Johnson then considers the welfare implications of a ban on loss leaders, which I only sketch out here. At the outset, Johnson shows that in the absence of consumer bias, more consumers would patronize the larger store. This is because the utility of shopping at the larger store increases because the consumer no longer underestimates the purchase likelihood of non-staples: In the presence of bias, consumers “underestimate . . . the value of goods in [the larger store’s] portfolio that [the smaller store] does not carry.” (p. 18.) Consequently, the utility from shopping at the larger store increases in the absence of bias. In this sense, larger stores suffer from under-patronage in the presence of consumer bias.

Thus, in the presence of consumer bias, a ban on loss leaders would harm consumers: “[C]onsumers underestimate their needs and tend to have unsatisfied demand [in patronizing the smaller store]—some would be better off if they instead shopped at the larger firm, even though it is less convenient. A ban on loss leading, by constraining the larger firm’s ability to attract consumers, leaves more needs ultimately unsatisfied and so reduces welfare.” (p. 19.)

Closing Observations

I really enjoyed reading this paper. First, it does incorporate one component of behavioral economics, bounded rationality. Second, as applied to loss leaders, the model is both simple and elegant. Third, and related, the conclusions are powerful: loss-leading by larger stores serves the interests of consumers with biased purchasing probabilities because it encourages them to visit the larger stores and so satisfy demands whose ex ante significance falls below their ex post sig-

¹⁴ Having said that, Johnson is quick to point out that price predation cannot be the explanation of the loss-leader practice since loss leaders are a persistent phenomenon, ruling out any recoupment. (p. 14.)

nificance. There is no anticompetitive motive by the larger firm in adopting a loss leader strategy. Indeed, in his model, consumers are better off by visiting the larger stores. It is difficult to predict the extent to which behavioral economics will influence antitrust policy, but papers like this certainly advance that cause.

To be sure, there are some assumptions that might give one pause. For example, in this model, consumers do not learn and so they do not revise their probability that they will purchase non-staples—the accuracy ratio does not increase. Having said that, it may be that bounded rationality limits the efficacy of updating.

In addition, the model assumes that consumers only shop at one store. But Johnson notes that “conclusions about the social effects of loss leading depend on whether the small firms are boutiques providing niche or high-end products that offer a large quality advantage, or rather small firms are more convenient for some customers but carry products that tend to be no higher quality than those of the large firm.” (p. 20.)¹⁵

Although the paper can be technical, much of the very technical material is in an appendix and Johnson spends time providing very good intuitive explanations for his results driven by the modeling. Thus, the paper should be understandable to antitrust practitioners of every stripe.

—JRW

¹⁵ Johnson cites other research suggesting that where “banning loss leading helps consumers” is where the small firm is a boutique. (p. 19.)