Journal of Competition Law & Economics, 13(1), 1–44
doi:10.1093/joclec/nhx006
Advance Access publication 2 March 2017

THE FALLACIES OF PATENT-HOLDUP THEORY

Alexander Galetovic[*] *& Stephen Haber*[†]

ABSTRACT

Patent-holdup theory avers that the patent system threatens the rate of innovation in the U.S. economy, particularly in information technology industries that are heavily reliant on standard-essential patents. We show that arrays of empirical tests falsify the core predictions of the theory. We therefore examine the logic of patent-holdup theory. We show that patent-holdup theory conflates two mutually inconsistent economic mechanisms: holdup (the appropriation of a quasi rent) and the exercise of monopoly power (to set the market price to extract a monopoly rent). Moreover, three fallacies underpin patent-holdup theory: (1) that patent holdup is a straightforward variant of holdup as it is understood in transaction-cost economics; (2) that royalty stacking is holdup repeated multiple times on the same product; and (3) that standard-essential patents contribute little or no value to the markets they help create. These fallacies give rise to a theory that is logically inconsistent and incomplete, and that ignores economic fundamentals. The flaws in logic of patent-holdup theory, and its lack of fit with the evidence, suggest that a new theory about the mechanics and dynamics of SEP-intensive IT industries is called for, both as a matter of science and as a guide to antitrust and patent policies.

JEL: D21; D23; K11; K12; L1; O31; O34; O38

I. INTRODUCTION

Until the late 1980s, archaeologists maintained that the Ancient Maya were a peace-loving people whose elites were primarily concerned with the scholarly

[*] Professor, Department of Economics, Universidad de los Andes, Santiago, Chile. Email: alexander@galetovic.cl.

[†] Peter and Helen Bing Senior Fellow, Hoover Institution, A.A. and Jeanne Welch Milligan Professor, School of Humanities and Sciences, Stanford University. Email: haber@stanford.edu.
We thank Aala Abdelgadir, Richard Epstein, Adriane Fresh, Wes Hartmann, Jordan Horrillo, Scott Kieff, Ross Levine, Noel Maurer, Victor Menaldo, Josiah Ober, Nicolas Petit, Gregory Sidak, Richard Sousa, Paul Sniderman, Daniel Spulber, Richard Stark, Seth Werfel, William Summerhill, and two anonymous referees for comments. Research support was provided by the Working Group on Intellectual Property, Innovation, and Prosperity (IP^2) of the Hoover Institution at Stanford University, which Haber directs. IP^2 succeeded the Hoover Project on Commercializing Innovation (PCI). To ensure academic freedom and independence, both PCI and IP^2, along with all work associated with them, have been supported only by unrestricted gifts. All such work, including this article, reflects the independent views of the authors. Some major donors include Microsoft, Pfizer, and Qualcomm. Galetovic previously consulted for Ericsson on a related topic.

study of astronomy and mathematics. They believed this conjecture despite the fact that one could not walk through a Mayan ruin without tripping over immense stelae depicting grotesquely violent images of victorious warriors subjugating their captives.

The process by which archaeologists created a theory about peaceful forest dwellers in the face of self-evident facts to the contrary is a testament to the power of fundamental fallacies. Their first fallacy was the idea that the inscriptions on the stelae were different from other glyph-based writing systems: instead of being a mix of whole words and phonetic sounds, as is the case with Egyptian hieroglyphic and cuneiform, archaeologists and epigraphers maintained that each symbol in the Mayan script represented an entire word or concept. That fallacy led them into a second fallacy: they maintained that non-calendrical Mayan hieroglyphs were indecipherable. That fallacy allowed the emergence of a third fallacy: because the theory of peace-loving forest people could not be tested against written evidence, the gruesome images on the stelae could be explained away as depicting mythical gods, not actual people. Thus, the archaeologists arrived at the false conclusion that the Maya were peaceful folk.

The fact that it took four decades for these fallacies to be overturned, one by one, is a testament to the reluctance of scholars to reject fashionable theories.[1] In point of fact, a Russian epigrapher had figured out the principles of Mayan translation in 1952, but it took two decades for American scholars to accept that his theory of Mayan writing was correct and theirs was wrong. It then took another decade for enough monumental inscriptions to be translated to convince archaeologists that the stelae did not depict mythical gods, but instead told the political history of Mayan kings—their birth, military conquests, and death. It took still another decade before a consensus emerged that the evidence that had been right in front of archaeologists all along contradicted their theory.

[1] The history of how Mayanists were able to maintain their fallacies for decades on end is also a testament to the influence of scholarly gatekeepers, most particularly Sir J. Eric Thompson of the Carnegie Institution and his acolytes. Their influence is laid bare by the fact that a 16th century Spanish Bishop, Diego de Landa, had worked with Mayan informants to transliterate the Spanish alphabet into corresponding Mayan glyphs, and that transliteration had been known to scholars since an abridged copy was unearthed and published in 1862. It was left to a Russian epigrapher, Yuri Knorozov, cut off from Western academia by the Cold War, to decipher that the "Landa Alphabet" was a syllabary, and thus Mayan symbols represented a mix of phonetic sounds and entire words. Knorozov's work was, however, dismissed and ridiculed by Thompson and his followers. Tatiana Proskouriakoff, a Russian émigrée architect, showed in the 1960s that Knorozov's translation system actually worked and the stelae were histories of Mayan elites. In the 1970s, some western scholars, particularly Michael D. Coe and David H. Kelly, began to build upon Knorozov's and Proskouriakoff's work and to challenge Thompson, who died in 1975. By the late 1980s, enough inscriptions had been translated that scholars accepted the view that the stelae documented the royal lineages and military conquests of Mayan kings. *See* MICHAEL D. COE, BREAKING THE MAYA CODE (Thames & Hudson 1992).

It would be comforting if the only field ever led astray by fundamental fallacies was Mayan archaeology, but that is hardly the case. Faulty premises often lead researchers toward conclusions that do not fit the facts—so much so that Nobel Laureate Richard Feynman made it the subject of his famous commencement address at the California Institute of Technology, in which he stressed the importance of bending over backward to do every test that might falsify a theory.[2]

A. An Influential Theory

Our concern here is with how fundamental fallacies gave rise to patent-holdup theory, which has guided antitrust and competition authorities around the world for nearly two decades. In the early 2000s, legal academics and antitrust economists asked an important question: does a decentralized system of technology development, in which complex, interoperable information technology (IT) products rely on standard-essential patents (SEPs) owned by many firms, allow SEP owners to "hold up" manufacturers, thereby stifling innovation and hurting consumers in the form of higher prices and lower-quality products?

The answer—patent-holdup theory—consists of five nested claims. First, that patent owners can systematically overcharge manufacturers for licenses to their patents through the economic mechanism of holdup—the opportunistic appropriation of a downstream firm's quasi rents (revenues in excess of short-run costs). Second, that when there are multiple patent holders, each practicing holdup on a downstream firm, cumulative patent royalty rates become astronomically high—a phenomenon patent-holdup theorists termed "royalty stacking." Third, that the holdup problem is exacerbated when patented technologies are included in the industry standards necessary to make IT products interoperable and compatible. Fourth, that patent holdup, royalty stacking, and the inclusion of patented technologies in industry standards are strangling innovation, most particularly in SEP-intensive IT products. Fifth, that the government must intervene to solve this problem; the market, left on its own, will fail.

Carl Shapiro's seminal article provides a clear statement of the threat posed by patent holdup to innovation:

> The holdup problem is worst in industries where hundreds if not thousands of patents, some already issued, others pending, can potentially read on a given product. In these industries, the danger that a manufacturer will step on a land mine is all too real. The result will be that some companies avoid the mine field altogether, that is, refrain from introducing certain products for fear of holdup.[3]

[2] Richard P. Feynman, *Cargo Cult Science*, 37 ENGINEERING & SCI., June 1974, at 10.

[3] Carl Shapiro, *Navigating the Patent Thicket: Cross Licenses, Patent Pools, and Standard Setting*, *in* 1 INNOVATION POLICY AND THE ECONOMY 121, 126 (Adam B. Jaffe, Josh Lerner & Scott Stern eds., MIT Press 2001).

He clearly articulates the need for a public policy intervention: "I submit that this holdup problem is very real today, and that both patent and antitrust policymakers should regard holdup as a problem of first order significance in the years ahead."[4]

The claim that patent holdup is common and is a threat to innovation can be found in any number of scholarly articles. Joseph Farrell, John Hayes, Carl Shapiro, and Theresa Sullivan state that "surprise hold-up may be largely a transfer, but anticipation of hold-up encourages a range of inefficient forms of self-protection, such as postponing or minimizing investment, or ensuring that standards use only antique technology."[5] Mark Lemley and Carl Shapiro concur:

> In the long run, if products are expected to be subject to some degree of holdup, the firm may not find it worth incurring the costs necessary to develop, manufacture, and sell the product. Assertions based on the shut-down condition that royalty stacking is somehow a minor problem or that royalty stacking cannot stifle innovation or hinder the market penetration of products that have been developed are simply unfounded.[6]

Most recently, Fiona Scott Morton and Carl Shapiro warn that patent holdup and its related mechanisms threaten the Internet of Things (IoT), and suggest the need for antitrust intervention:

> Failure to prevent Patent Holdup relating to tomorrow's information technology and communications standards is likely to cause significant social welfare loss in the years ahead. If new and more effective private solutions relating to standard setting do not emerge to promote innovation and protect consumers, antitrust enforcement is one of the only remaining remedies that seems feasible.[7]

Patent-holdup theory has also been influential among antitrust authorities around the world. Several Federal Trade Commission (FTC) reports[8] and a

[4] *Id.* at 125. Government intervention is also suggested in Joseph Farrell, John Hayes, Carl Shapiro & Theresa Sullivan, *Standard Setting, Patents and Hold-up*, 74 ANTITRUST L.J. 603 (2007); Mark A. Lemley, *Ten Things to Do About Patent Holdup of Standards and One Not To*, 48 B.C. L. REV. 149 (2007); Mark A. Lemley & Carl Shapiro, *Patent Holdup and Royalty Stacking*, 85 TEX. L. REV. 1991 (2007); Carl Shapiro, *Patent Reform: Aligning Reward and Contribution*, *in* 8 INNOVATION POLICY AND THE ECONOMY 111 (Adam Jaffe, Josh Lerner & Scott Stern eds., Univ. of Chicago Press 2008) [hereinafter Shapiro, *Patent Reform*]; Carl Shapiro, *Injunctions, Hold-up, and Patent Royalties*, 12 AM. L. & ECON. REV. 280 (2010); Fiona Scott Morton & Carl Shapiro, *Patent Assertions: Are We Any Closer to Aligning Reward to Contribution?*, *in* 16 INNOVATION POLICY AND THE ECONOMY 89 (Josh Lerner & Scott Stern eds., Univ. of Chicago Press 2016) [hereinafter Scott Morton & Shapiro, *Patent Assertions*].

[5] Farrell, Hayes, Shapiro & Sullivan, *supra* note 4, at 647.

[6] Lemley & Shapiro, *Patent Holdup and Royalty Stacking*, *supra* note 4, at 2012.

[7] Scott Morton & Shapiro, *Patent Assertions*, *supra* note 4, at 124. Other theorists, such as Jorge Contreras and Richard Gilbert, have argued that the same problems affect all products that include patents, not solely those that rely on SEPs. *See, e.g.*, Jorge L. Contreras & Richard J. Gilbert, *A Unified Framework for RAND and Other Reasonable Royalties*, 30 BERKELEY TECH. L.J. 1451 (2015).

[8] Federal Trade Commission, To Promote Innovation: The Proper Balance of Competition and Patent Law and Policy (2003) [hereinafter FTC, To Promote Innovation]; Federal Trade

joint Department of Justice (DOJ) and FTC report[9] discuss the threat to innovation posed by patent holdup and royalty stacking, citing the academic literature. For example, one FTC report states:

> Unless downstream actors—whether innovators or manufacturers—can mitigate the problem [of patent holdup], they may have to choose between the risk of being sued for infringement after they sink costs into invention or production, or dropping innovative or productive efforts altogether. Either option can injure economic welfare.[10]

These views are echoed by agency heads, such as the former chair of the FTC, the former Deputy Attorney General for Antitrust, and the European Competition Commissioner, who signal that they are willing to take action about the problem.[11]

Patent-holdup theory is also mentioned in amicus briefs that argue that patent holdup is a common occurrence. For example, a 2006 brief file by 52 intellectual property professors submitted in support of the defendant in e*Bay, Inc. v. MercExchange, L.L.C.* states that:

Commission, The Evolving IP Marketplace: Aligning Patent Notice and Remedies with Competition (2011) [hereinafter FTC, The Evolving IP Marketplace].

[9] U.S. Department of Justice & Federal Trade Commission, Antitrust Enforcement and Intellectual Property Rights: Promoting Innovation and Competition (2007) [hereinafter DOJ & FTC, Antitrust Enforcement].

[10] FTC, To Promote Innovation, *supra* note 8, at 28.

[11] *See, e.g.*, Bill Baer, Assistant Attorney Gen., Antitrust Div., U.S. Dep't of Justice, Remarks as Prepared for Delivery at the 19th Annual International Bar Association Competition Conference, Reflections on the Role of Competition Agencies When Patents Become Essential (Sept. 11, 2015); Renata Hesse, Deputy Assistant Attorney Gen., Antitrust Div., U.S. Dep't of Justice, Remarks as Prepared for the ITU-T Patent Roundtable, Six "Small" Proposals for SSOs Before Lunch (Oct. 10, 2012); Renata Hesse, Deputy Assistant Attorney Gen., Antitrust Div., U.S. Dep't of Justice, Remarks Presented at the Global Competition Review: 2nd Annual Antitrust Law Leaders Forum, IP, Antitrust and Looking Back on the Last Four Years (Feb. 8, 2013); Renata Hesse, Deputy Assistant Attorney Gen., Antitrust Div., U.S. Dep't of Justice, Remarks Presented at the Washington State Bar Association, The Art of Persuasion: Competition Advocacy at the Intersection of Antitrust and Intellectual Property (Nov. 8, 2013); Renata Hesse, Deputy Assistant Attorney Gen., Antitrust Div., U.S. Dep't of Justice, Remarks as Prepared for the Global Competition Review: GCR Live IP & Antitrust USA Conference, A Year in the Life of the Joint DOJ-PTO Policy Statement on Remedies for F/RAND Encumbered Standards-Essential Patents (Mar. 25, 2014); Edith Ramirez, Chairwoman, FTC, 2014, Address at the 8th Annual Global Antitrust Enforcement Symposium, Standard-Essential Patents and Licensing: An Antitrust Enforcement Perspective (Sept. 10, 2014) [hereinafter Ramirez, SEPs and Licensing]; *see also* U.S. Department of Justice & U.S. Patent & Trademark Office, Policy Statement on Remedies for Standards-Essential Patents Subject to Voluntary F/RAND Commitments (2013) [hereinafter DOJ & USPTO, Policy Statement]; Margrethe Vestager, Comm'r for Competition, Eur. Comm'n, Address at the Chillin' Competition Conference, Protecting Consumers from Exploitation (Nov. 21, 2016). For a discussion of Canada and Japan, see Lisa Kimmel, *Injunctive Relief for Infringement of FRAND-Assured Standard-Essential Patents: Japan and Canada Propose New Antitrust Guidance*, CPI Antitrust Chron., Oct. 2015, at 1.

> [S]uch inappropriate "holdups" occur on a regular basis under the Federal Circuit's mandatory-injunction standard. Patentees can obtain revenue in excess of the value of their technology by threatening to enjoin products that are predominantly noninfringing and in which the defendant has made significant irreversible investments.[12]

It should therefore not be surprising that courts have been influenced by patent-holdup theory. For example, in *eBay*, Justice Kennedy's concurring opinion cites an FTC report that warns of the impact of patent holdup by firms that do not themselves practice their patents:

> An industry has developed in which firms use patents not as a basis for producing and selling goods but, instead, primarily for obtaining licensing fees. For these firms, an injunction, and the potentially serious sanctions arising from its violation, can be employed as a bargaining tool to charge exorbitant fees to companies that seek to buy licenses to practice the patent. When the patented invention is but a small component of the product the companies seek to produce and the threat of an injunction is employed simply for undue leverage in negotiations, legal damages may well be sufficient to compensate for the infringement and an injunction may not serve the public interest.[13]

The landmark Supreme Court *eBay* decision is not an outlier. Jonathan Barnett identifies thirty-seven federal court decisions that mention "patent holdup" or "royalty stacking."[14]

B. The Stelae That Contradicted the Theory

Like the theory of the peaceful Maya, patent-holdup theory had its own set of facts—stelae, as it were—that contradicted the theory. Patent-holdup theorists asserted that innovation in SEP-intensive IT products was under threat: excessive royalties were discouraging new firm entry and reinvestment by existing firms. They called particular attention to the threat to innovation in mobile telephones and personal computers, as well as in extensions of those products in the IoT.[15]

[12] Brief Amici Curiae of 52 Intellectual Property Professors in Support of Petitioners, eBay, Inc. v. MercExchange, L.L.C., 547 U.S. 388 (2006) (No. 05-130), 2006 WL 1785363.

[13] *eBay*, 547 U.S. at 396–97 (2006) (Kennedy, J., concurring) (citing FTC, To Promote Innovation, *supra* note 4, at 38–39). For a critical assessment of Justice Kennedy's concurrence, co-authored by a retired chief judge of the Federal Circuit, see Paul R. Michel & Michael J. Dowd, *Understanding the Errors of* eBay, 2 Criterion J. on Innovation 21 (2017).

[14] Jonathan Barnett, *Has the Academy Led Patent Law Astray?*, 32 Berkeley Tech. L.J. (forthcoming 2017). Barnett also identifies eight International Trade Commission proceedings, and four FTC proceedings that mention patent holdup or royalty stacking.

[15] For example, Lemley and Shapiro state: "In the information technology sector in particular, modern products such as microprocessors, cell phones, or memory devices can easily be covered by dozens or even hundreds of different patents. As a striking example, literally thousands of patents have been identified as essential to the proposed new standards for 3 G cellular telephone systems." Lemley & Shapiro, *Patent Holdup and Royalty Stacking*, *supra* note 4, at 1992. Their case studies focus on 3 G cellular technologies, Wi-Fi 802.11 technologies, DVD media, the MP3 music format, and RFID chips. *Id.* at 2025–29. With the

Economists measure rates of innovation by examining relative rates of change of quality-adjusted prices,[16] and one can download the publicly available, product-by-product, quality-adjusted price data compiled by the Bureau of Labor Statistics in order to carry out an analysis of innovation rates across products and within products over time.[17] An analysis of that data shows that from 1997 to 2013, rates of innovation in phone equipment (which includes low-tech items such as fax machines and landline phones, as well as wireless phones) was 10 percent per annum faster than the economy-wide average. The data show that the rate of innovation in portable and laptop computers was faster still—31 percent per annum faster than the economy-wide average. Similar rates of innovation are observed in other SEP-intensive IT products such as video equipment, audio equipment, desktop computers, and televisions. Furthermore, rates of innovation in SEP-intensive IT products have not slowed over time relative to the rates of innovation in similar, non-SEP-intensive IT products.[18] For example, the rate of innovation

exception of RFID chips, all of the other technologies they study are necessary for the interoperability and compatibility of a smartphone. Personal computers require all of these technologies, except for 3 G mobility and RFID. Farrell, Hayes, Shapiro, and Sullivan also call attention to the potential problem in smartphones and personal computers, motivating their article with seven cases, all of which are applicable to these product lines. Farrell, Hayes, Shapiro & Sullivan, *supra* note 4. Scott Morton and Shapiro point to the IoT as an area of concern. The IoT is an extension of smartphone and personal computer technologies to other devices. *See* Scott Morton & Shapiro, *Patent Assertions*, *supra* note 4.

[16] This approach to measuring innovation was pioneered by several giants in the economics profession, most particularly Dale Jorgenson and Zvi Griliches, and has since spurred a broad and distinguished literature on the measurement and sources of productivity growth. *See* Dale W. Jorgenson & Zvi Griliches, *The Explanation of Productivity Change*, 34 REV. ECON. STUD. 249 (1967); *see, e.g.*, Dale W. Jorgenson, *Accounting for Growth in the Information Age*, PRODUCTIVITY AND CYCLICALITY IN SEMICONDUCTORS: TRENDS, IMPLICATIONS, AND QUESTIONS (Dale W. Jorgenson & Charles W. Wessner eds., National Research Council 2004); William D. Nordhaus, *Two Centuries of Productivity Growth in Computing*, 67 J. ECON. HIST. 128 (2007); Kenneth Flamm, Economic Benefits from Technological Innovation in Microelectronics (Nov. 2010) (unpublished manuscript), https://www.nsf.gov/sbe/sosp/tech/flamm.pdf. For an explanation of the methods that the Bureau of Labor Statistics uses to adjust prices for quality changes, see Brent R. Moulton & Karin E. Moses, *Addressing the Quality Change Issue in the Consumer Price Index*, 1997 BROOKINGS PAPERS ON ECON. ACTIVITY 1, 305; Mark Bils, *Do Higher Prices for New Goods Reflect Quality Growth or Inflation?*, 124 Q.J. ECON. 637 (2009).

[17] The relationship between rates of innovation and prices is alluded to in the literature on patent holdup. *See, e.g.*, Shapiro, *Navigating the Patent Thicket: Cross Licenses, Patent Pools, and Standard Setting*, *supra* note 3, at 126; Lemley, *Ten Things to Do About Patent Holdup of Standards and One Not To*, *supra* note 4, at 152; Farrell, Hayes, Shapiro & Sullivan, *supra* note 4, at 608; DOJ & FTC, ANTITRUST ENFORCEMENT, *supra* note 9, at 35; FTC, THE EVOLVING IP Marketplace, *supra* note 8, at 4, 22 28, 30, 78, 140, 145, 160.

[18] *See* Alexander Galetovic, Stephen Haber & Ross Levine, *An Empirical Examination of Patent Holdup*, 11 J. COMPETITION L. & ECON. 549 (2015). No matter how the authors treat the data—including the use of a difference in differences estimator on de-trended data in order to control for heterogeneous underlying rates of innovation across products—that study cannot produce a result that is consistent with the hypothesis that relative rates of innovation in SEP-intensive IT products have been slower than non-SEP-intensive IT products. Alexander Galetovic and Kirti

in SEP-intensive laptop computers compared with non-SEP-intensive mainframe computers shows that SEP intensity was associated with faster innovation.[19]

There are other hallmarks of innovation beyond falling quality-adjusted relative prices: one would expect to see rapidly increasing output even in the face of falling prices; and, because innovation is typically characterized by Schumpeterian creative destruction, one would also expect to see high levels of firm entry and exit. This is precisely what researchers do see when they examine data on the canonical case of the mobile phone industry. Between 1994 and 2013, the number of SEP holders increased from 2 to 128. Patent-holdup theory would predict that this increase should have dramatically slowed the rate of innovation. That prediction did not obtain in reality, however. Prices of mobile devices dropped very fast, while output grew sixty-two-fold. During this same period, there was rapid entry of new firms into the manufacture of phones and tablets—so much so that industrial concentration, measured with the number of devices sold, actually fell in this industry over time.[20]

According to patent-holdup theory, excessive patent royalties cause slow rates of innovation. As an empirical matter, the aggregate royalties paid by licensees in any industry can be estimated on the basis of the SEC 10-K and 40-F filings of the patent-licensing firms.[21] The data on the canonical case of mobile phones shows that the cumulative royalty yield from the twenty-one largest patent licensors in the mobile phone value chain was only 3.3 percent of a mobile phone's average sales price in 2015. That ratio has been fairly

Gupta obtain qualitatively similar results by looking across and within generations of mobile phones and tablets. Alexander Galetovic & Kirti Gupta, *Royalty Stacking and Standard Essential Patents: Theory and Evidence from the World Wireless Industry* (Hoover IP2 Working Paper Series No. 15012, 2016). Each generation of phone and tablet (for example, 2 G, 2.5 G, 3 G, 3.5 G, and 4 G) can be thought of as capturing an improvement in quality. Galetovic and Gupta find that the introduction of every successive generation of phone and tablet came with a lower average selling price than the previous generation—even though each successor generation was a superior product. They also show that within each generation, prices fell at rates between 10 and 20 percent per year. Indeed, if innovation really was stagnating, then there should not have been successive generations of phones or tablets. Patent holdup and its related mechanisms of royalty stacking and market power conferred by standards should have choked off the incentives to invest in R&D.

[19] Mainframes and laptops are both digital products that are subject to Moore's Law. But mainframes are not as SEP intensive because they do not need to be interoperable and compatible. They are special-purpose machines that use proprietary architectures and run proprietary software.

[20] *See* Galetovic & Gupta, *supra* note 18; *see also* Kirti Gupta, *Technology Standards and Competition in the Mobile Wireless Industry*, 22 GEO. MASON L. REV. 865 (2015); Keith Mallinson, *Don't Fix What Isn't Broken: The Extraordinary Record of Innovations and Success in the Cellular Industry Under Existing Licensing Practices*, 23 GEO. MASON L. REV. 967 (2016) [hereinafter Mallinson, *Don't Fix What Isn't Broken*].

[21] As a matter of accounting, every dollar paid by a manufacturer for a patent license must appear on the revenue statements of the entities that own the patents.

stable since at least far back as 2007.[22] Researchers have parameterized royalty-stacking models from the patent-holdup literature using actual price and quantity data, and have discovered that the royalty yield predicted by the models is more than twenty times higher than the actual royalty yield and about four-fifths of the price of a smartphone. They have also found that no individual patent licensor earns an individual royalty consistent with the hypothesis that it operated as a monopolist.[23]

The facts of fast and continuous innovation in the mobile phone industry—one of the stelae of patent-holdup theory—are evident to anyone with a smartphone in their pocket. Three decades ago, a mobile phone cost the current equivalent of $10,000, was the size of a brick, weighed a kilo, and enabled its user to make a half-hour call before going dead. Today, a smartphone has more computational power than the supercomputers that guided the Apollo missions to the moon, allows a user to produce and share data, video, and audio files with anyone on the planet, costs an average of $300—and also happens to make a phone call.

At the same time that there are self-evident stelae contradicting patent-holdup theory, there is no positive evidence in support of its core predictions. Damien Geradin and Miguel Rato,[24] Damien Geradin, Anne Layne-Farrar and Jorge Padilla,[25] Vincenzo Denicolò, Damien Geradin, Anne Layne-Farrar, and Jorge Padilla,[26] Richard Epstein, F. Scott Kieff, and

[22] *See* Alexander Galetovic, Stephen Haber & Lew Zaretzki, *A New Dataset on Mobile Phone Patent License Royalties: September 2016 Update* (Hoover IP2 Working Paper Series No. 16011, 2016). Galetovic, Haber, and Zaretski build upon work by Keith Mallinson, who estimated an upper bound of the royalty yield on SEPs of about 5.5 percent. Keith Mallinson, *Cumulative Mobile-SEP Royalty Payments No More Than Around 5% of Mobile Handset Revenues*, IP FINANCE (Aug. 19, 2015), http://www.ip.finance/2015/08/cumulative-mobile-sep-royalty-payments.html. The differences in results are largely the result of the fact that Galetovic, Haber, and Zaretzki are able to make more fine-grained estimates of royalties for patent assertion entities and patent pools than Mallinson. Mallinson, *Don't Fix What Isn't Broken*, *supra* note 20. J. Gregory Sidak builds upon Mallinson as well, but departs from him in that he also includes payments in kind and estimates of the value of cross-licenses. He reaches an upper bound of between 4 and 5 percent. J. Gregory Sidak, *What Aggregate Royalty Do Manufacturers of Mobile Phones Pay to License Standard-Essential Patents?*, 1 CRITERION J. ON INNOVATION 701 (2016).

[23] *See* Alexander Galetovic, Stephen Haber & Lew Zaretzki, *Is There an Anti-Commons Tragedy in the Smartphone Industry?*, 32 BERKELEY TECH. L.J. (forthcoming 2017).

[24] Damien Geradin & Miguel Rato, Can Standard-Setting Lead to Exploitative Abuse? A Dissonant View on Patent Hold-Up, Royalty Stacking and the Meaning of Frand, 3 EUR. COMPETITION J. 101 (2007).

[25] 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).

[26] Vincenzo Denicolò, Damien Geradin, Anne Layne-Farrar & A. Jorge Padilla, *Revisiting Injunctive Relief: Interpreting* eBay *in High-Tech Industries with Non-Practicing Patent Holders*, 4 J. COMPETITION L. & ECON. 571 (2008) [hereinafter Denicolò, Geradin, Layne-Farrar & Padilla, *Revisiting Injunctive Relief*].

Daniel Spulber,[27] Kirti Gupta,[28] Anne Layne-Farrar,[29] J. Gregory Sidak,[30] and Edward Egan and David Teece[31] review the literature on patent holdup, patent thickets, and royalty stacking.[32] All of these studies reach the same general conclusion, which is perhaps best summed up by Layne-Farrar: "Certainly the theories have been developed, but the empirical support is still lacking. Despite the fifteen years that proponents of the theories have had to amass evidence, the empirical studies conducted thus far have not shown that holdup or royalty stacking is a common problem in practice."[33]

[27] Richard A. Epstein, F. Scott Kieff & Daniel Spulber, *The FTC, IP, and SSOS: Government Hold-up Replacing Private Coordination*, 8 J. COMPETITION L. & ECON. 1 (2012).

[28] Kirti Gupta, *The Patent Policy Debate in the High-Tech World*, 9 J. COMPETITION L. & ECON. 827 (2013).

[29] Anne Layne-Farrar, Patent Holdup and Royalty Stacking Theory and Evidence: Where Do We Stand After 15 Years of History? (OECD 2014).

[30] J. Gregory Sidak, *Apportionment, FRAND Royalties, and Comparable Licenses After* Ericsson v. D-Link, 2016 U. ILL. L. REV. 1809.

[31] Edward J. Egan & David Teece, *Untangling the Patent Thicket Literature* (Tusher Ctr. for Mgmt. of Intellectual Capital, Working Paper No. 7, 2015).

[32] Some of these studies call into question the facts presented in the extant literature. Denicolò, Geradin, Layne-Farrar, and Padilla, for example, question the two cases, *Rambus* and *RIM-BlackBerry*, that Lemley and Shapiro present as evidence of patent holdup, and state that in one case the facts are not correctly stated and in the other the necessary economic analysis is not carried out. Denicolò, Geradin, Layne-Farrar & Padilla, *Revisiting Injunctive Relief*, *supra* note 26, at 597–99; Lemley & Shapiro, *Patent Holdup and Royalty Stacking*, *supra* note 4. Epstein, Kieff, and Spulber criticize an influential FTC report: FTC, THE EVOLVING IP MARKETPLACE, *supra* note 8. They state that the report "does not offer any quantitative estimate of value destroying breakdowns on the present system, nor does it offer any empirical basis to conclude that such breakdowns are of a frequency and magnitude that could justify radical change to a system that has enabled numerous successfully implemented standards." Epstein, Kieff & Spulber, *supra* note 27, at 13.

[33] Layne-Farrar, *supra* note 29, at 2. Some studies count the number of patents or patent holders necessary to produce a product. *See, e.g.*, Jorge L. Contreras, *Standards, Royalty Stacking, and Collective Action*, CPI ANTITRUST CHRON., Mar. 2015, at 1, 2. Others cite to the literature on the measurement of fragmentation of patent ownership—so-called patent thickets—as evidence of royalty stacking. *See, e.g.*, Lemley & Shapiro, *Patent Holdup and Royalty Stacking*, *supra* note 4. Showing that there are many patent holders in an industry or that their ownership is fragmented is, however, only the first step in making the case that royalty stacking is holding back innovation. As a second step, a researcher must show that those patent holders have market power because they can set prices. Then, as Daniel Spulber explains, as a third step, a researcher needs to show that those patent holders independently post a linear, per-unit royalty. Daniel F. Spulber, *Complementary Monopolies and Bargaining* (Northwestern Law & Econ. Research Paper No. 16-10, 2016). Finally, as a fourth step, a researcher must show that, as the number of such price-setting patent holders increases, the equilibrium price of the final good increases, while output falls. To our knowledge, there is only one article, by Iain Cockburn, Megan Macgarvie, and Elizabeth Muller (studying a sample of German firms) that comes close to accomplishing these steps. *See* Iain M. Cockburn, Megan J. MacGarvie & Elisabeth Muller, *Patent Thickets, Licensing, and Innovative Performance*, 19 INDUS. & CORP. CHANGE 899 (2010). This article, however, actually finds evidence against the royalty-stacking hypothesis. Indeed, the firm in the sample with the highest cumulative royalty spent only 2.12 percent of its sales on patent licenses. That figure, as we explain in Galetovic, Haber & Zaretzki, *Is There an Anti-Commons Tragedy in the Smartphone Industry*, *supra* note 23, is at least one order of magnitude

C. Three Fundamental Fallacies and Their Origin

When theory and evidence disagree, there is either something wrong with the theory or something wrong with the evidence. We think that there is something wrong with the theory.

Patent-holdup theory conflates two different economic mechanisms: holdup and market power. Holdup means that one firm appropriates another firm's quasi rent—its revenues minus its short-run costs—through opportunistic behavior. A firm that is being held up, by definition, does not generate enough revenue to cover its long-run costs. Therefore, the firm will not reinvest once its capital wears out. This is not a long-run equilibrium. Market power, by contrast, means that a firm can set prices such that it appropriates a monopoly rent from a market. The exercise of market power can be a long-run equilibrium, because the downstream firms will cover their long-run costs and continue to reinvest as their capital equipment wears out.[34] Thus, holdup and the exercise of market power are two different, mutually inconsistent economic mechanisms. One cannot simultaneously have a long-run equilibrium and not have a long-run equilibrium.

The conflation of holdup and market power leads to three fallacies that underpin patent-holdup theory. Once the mechanics of holdup are loosened from their moorings in economic theory, it becomes possible to simultaneously claim that patent holdup is a variant of holdup as it is understood in mainstream economics and define it in ways that are inconsistent with the meaning of holdup as it is understood in mainstream economics. Patent holdup elides key assumptions of the standard theory and transforms necessary conditions for holdup into sufficient conditions for holdup. The implications are fundamental. In the established theory, firms—working together—will make structural, contractual, and behavioral adaptations in order to prevent holdup, thereby sustaining trade and investment in equilibrium. In patent-holdup theory, by contrast, firms cannot adapt and solve the problem wrought by opportunistic renegotiation of a contract, because the game begins after the R&D is completed and manufacturers invest. Adaptations to prevent holdup are ruled out by construction, and market failure is inevitable.

The conflation of holdup and market power leads to a second fallacy. Patent-holdup theory claims that the same manufacturing firms can be held up many times over, resulting in a phenomenon called royalty stacking. In point of fact, however, holdup cannot occur many times over to the same firm. A firm's quasi rents (the difference between its revenues and its

smaller than one would expect in an industry characterized by royalty stacking. The average observed royalty burden in the sample was only 0.054 percent of sales, which is to say that it is at least three orders of magnitude smaller than one would expect in an industry plagued by royalty stacking.

[34] On the definition of quasi rent and monopoly rent, see Roger G. Noll, *Buyer Power and Economic Policy*, 72 ANTITRUST L.J. 589, 593 (2005).

short-run costs) can be extracted only once. Any attempt to extract more revenues would cause the firm to shut down. Royalty stacking, by contrast, is about the exercise of market power by multiple input suppliers to downstream firms. Although this multiplicity of input suppliers might be an inefficient organization of a market, it nonetheless can be a long-run equilibrium, unlike holdup.

To claim that market power is being exercised, one needs to identify its source. In royalty stacking, the source is the patents themselves. A patent confers a temporary, limited property right that might confer some market power—and does so by design. Thus, in order to claim that there is a public policy problem, one needs to claim that the patents in question confer market power in excess of that which is conferred by the patent grants themselves. What could the source of that excess market power be? According to the theory, patent-holding firms are able to appropriate more than their incremental contribution to a product's value by virtue of the fact that their technologies have been made part of a standard. The users of the technology are locked into that standard and consequently can be subjected to patent holdup.

The conflation of holdup and market power leads to a third fallacy: patented technologies that are part of an industry standard add little or no value to the markets that they help to create. There are two problems with this fallacy—one theoretical and the other empirical. The theoretical problem is, as Nobel Prize winner Kenneth Arrow showed in 1962, that when an innovation is "drastic" (that is, much better than the alternatives on offer) a profit-maximizing monopoly will charge *less* than the technology's incremental value. The empirical problem is that the whole point of standard development organizations (SDOs) in IT industries is to make large technological jumps at a fast pace, so that manufacturers may produce superior products that consumers will adopt enthusiastically, thereby increasing the revenues of all the industry stakeholders.[35] They are not in the business of small incremental improvements; they are in the business of creating drastic innovations.

D. The Goals of This Article

Our purpose is to analyze and dissect patent-holdup theory in order to show how it is based on a set of fallacies. These fallacies are not merely a matter

[35] A quintessential example is the successive changes in mobile phone standards that have allowed end-user data rates to increase more than 1000-fold since the 1990s, thereby allowing clunky 1 G cellphones to evolve into the mobile, high-speed computers, entertainment systems, medical monitoring devices, and communications platforms that are called smartphones. That process of standard setting, R&D, and patent licensing created a product so highly valued by consumers that there are now more mobile device connections than there are adult human beings on the planet. It also produced nearly $2.8 trillion in revenues for device manufacturers from 2007 through 2015. For information on data speeds by technology generation and market penetration, see Mallinson, *Don't Fix What Isn't Broken*, *supra* note 20.

of semantics. Microeconomics is a science of mechanics. Terms such as "holdup," "market power," and "monopoly" have precise meanings; they describe particular mechanisms and states of the world. Conflating them with one another gives rise to misunderstandings, which then multiply as the theory expands. The end result can be a theory whose predictions are starkly at odds with the empirical facts.

Our purpose is neither to advance an alternative theory, nor to evaluate whether there could be improvements in the contracting environment in SEP-intensive IT industries. It is almost always true that participants in a market—even participants that are thriving—can point to particular features of their contracting environment that they wish were different. We therefore do not take positions on whether injunctions should be granted to SEP holders or whether any particular SDO might reform its practices. Rather, our goal is to make a more fundamental point: that any set of policies or reforms whose purpose is to improve a contracting environment should not be based on patent-holdup theory.

II. THE THEORY OF TRANSACTION-COST HOLDUP

Microeconomics is concerned with how rational firms, households, and individuals adapt to changes in tastes, endowments, technology, and government policies. The terms it has developed have precise meanings; they are not mere labels or names, but terms of art.

In this part we show that holdup and market power are terms of art that denote two different and mutually inconsistent economic mechanisms. To hold up means to appropriate a firm's quasi rent—its total revenues minus its short-run costs—through opportunistic behavior. Because firms that are held up do not cover their long-run costs, firms that have been held up will not reinvest once their capital wears out. Hence, holdup cannot be a long-run equilibrium. If all firms in a market are being held up, the market will die. Market power, by contrast, means that a firm has the ability to set prices in a market such that it can extract a monopoly rent—the excess of total market revenues over long-run costs.[36] The exercise of market power can be a long-run equilibrium; downstream firms can still cover their long-run costs, and thus continue to reinvest as their capital equipment wears out. The vast difference between the two mechanics is captured by a simple fact: it is not necessary for a firm to have market power in order to hold up another firm; a firm can be held up even in a perfectly competitive input market.

These insights require nothing more than basic economics. We discuss them, however, because there is some confusion in the patent-holdup literature about the meaning of these concepts. Basic as the concepts of holdup

[36] *See* Noll, *supra* note 34, at 593.

and market power are, their conflation in the literature requires that we return to basics.

A. What Is Holdup?

Let us illustrate the economic concept of holdup by way of an example. Your local coffee bars operate in a competitive market. Each charges the market price for an espresso, and earns a normal rate of return on its capital. The market price allows coffee bars to cover: (1) variable costs (for example, barista wages, coffee, and supplies), which are incurred only if a bar is open for business; (2) fixed costs (for example, rent, heating, and telephone), which are incurred whether a bar is open or not; (3) the opportunity cost of capital (the amount an owner would have earned had she invested in U.S. treasury bonds, rather than bought a commercial grade espresso machine and upscale decorations); and (4) the value of the risk that the owner took when investing in the coffee bar.

Coffee bars buy most, if not all, of their inputs in competitive markets. In what follows, we will study two different market equilibria when one key input—real estate—is alternatively offered competitively or by a landlord who owns all real estate. We then ask, can a monopoly landlord simultaneously hold up the coffee shop and exploit monopoly power? The answer, as we shall show below, is no; they are mutually inconsistent economic mechanisms.

1. Equilibrium in a Competitive Market

We begin by assuming a competitive real estate market. Figure 1 shows the equilibrium in the market for espressos. The long-run supply curve equals the long-run average (and marginal) cost of producing espressos, c_E, which equals the sum of coffee bars' per-unit variable costs, fixed costs (including the competitive rent paid every month for the real estate), opportunity cost of capital, and the risk premium for being in the coffee bar business.[37] With competitive espresso and real estate markets (which we denote by the capital letter C), the equilibrium price, p_C, and quantity of espressos produced, X_C, are determined by the intersection of demand and the long-run supply curve for espressos.

There are two well-known properties of this equilibrium. First, coffee bar owners reinvest when their equipment wears out because they are covering their long-run costs. Second, the economic surplus created by the sale of

[37] Following standard practice, assume that in this perfectly competitive industry there are many coffee bars and that each has the same textbook U-shaped long-run average cost curve. In a long-run equilibrium, all coffee bars operate at their efficient scale, so that price equals both minimum long-run average cost and long-run marginal cost. The long-run market supply curve is the envelope of all coffee bars' long-run average costs at their efficient scale. Hence the aggregate long-run supply curve is flat and is equal to the minimum long-run average cost.

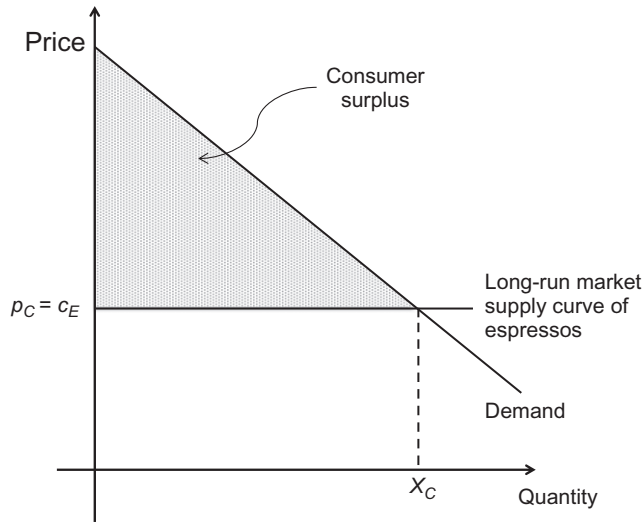

Figure 1. The long-run equilibrium of the coffee bar market with a competitive real estate market

espressos is the area between the demand curve and the supply curve—the shaded triangle in Figure 1. Because both the real estate and coffee bar markets are competitive, this is the maximum possible surplus that can be created in the espresso market, and consumers keep all of that surplus.

We cannot stress more strongly that the total amount of surplus available in any market is bounded by the demand curve. Surplus comes not from the inputs to production, but from consumer demand for the final product—in this case, an espresso.

2. *Equilibrium with a Monopoly Input Provider*

Imagine now that your local coffee bar is in a city with only one monopolist landlord. The landlord understands that all coffee bars need his real estate and that economic value is created by consumers who value espressos.[38] The coffee bar owners understand this too. The landlord therefore exploits his market power over espresso drinkers by charging, in addition to the competitive rent, an additional fee paid by coffee bar owners per espresso, which we

[38] Joseph Spengler showed that an upstream monopoly selling an input to a competitive downstream industry that uses it in fixed proportions can fully exploit monopoly power by appropriately choosing the price of the input. Joseph J. Spengler, *Vertical Integration and Antitrust Policy*, 58 J. POL. ECON. 347 (1950). Our aim here, however, is to stress that whatever market power rent the upstream monopoly makes, it comes from the demand for the final good, in this case espressos. For this reason, we will ignore the complications that arise when downstream firms can substitute other inputs for real estate.

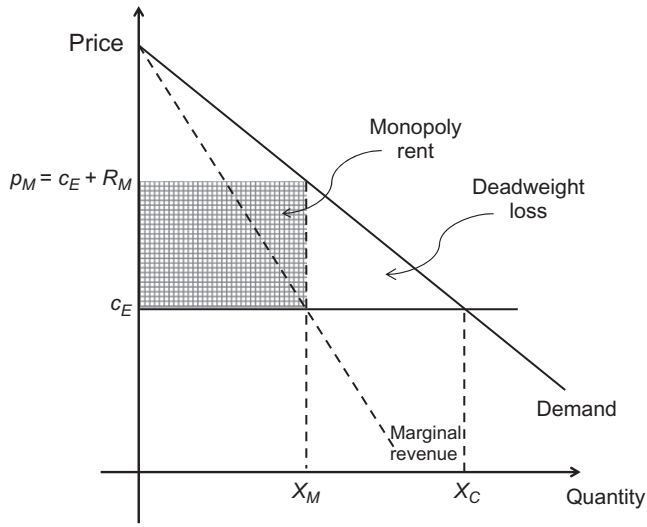

Figure 2. The long-run equilibrium of the coffee bar market with a single landlord

denote as R, to earn what economists call a monopoly rent.[39] He obtains this monopoly rent through the coffee bar owners.

Standard economic theory says that, as Figure 2 shows, the landlord will increase R until the marginal revenue from selling espressos is equal to coffee bars' average long run-costs, c_E.[40] Because coffee bars compete, they will pass R through to consumers. In equilibrium, the price of an espresso will be equal to $p_M = c_E + R_M$ (where we use the subscript M to denote the monopoly equilibrium). Consequently, compared with a competitive real estate market, coffee bar owners sell fewer espressos; quantity falls from X_C to X_M.

None of this is a surprise to coffee bar owners or would-be coffee bar owners; everyone knows that their pricing must reflect the per-espresso fee R. Consequently, all coffee bars cover their long-run costs and coffee bar owners reinvest when their equipment wears out. This market, like the market depicted in Figure 1, is in long-run equilibrium.

Note that once we introduce a real estate monopoly, the realized economic surplus created by espresso sales shrinks. The landlord's fee per espresso, R, increases the price of an espresso above the long-run marginal cost of producing it. Thus, fewer espressos are sold, and the total economic surplus is smaller. This deadweight loss is indicated in Figure 2 by the triangle. Conceptually, the deadweight loss triangle represents the sales of espresso that are not made because the price is too high.

[39] For a definition of monopoly rent, see Noll, *supra* note 34, at 593. Note that shopping malls typically charge shop owners a percentage of sales for their real estate (even though normally they are not monopolies).

[40] Note that $\frac{p_M - c_E}{p_M} = \frac{R_M}{p_M}$ is the well-known Lerner margin.

Note also that there is a different distribution of the surplus created by the espresso market once there is a monopolist landlord. When there was a competitive market for retail space, the entire surplus from the espresso market was captured by consumers, but now consumers and the landlord share the smaller surplus.

3. Holdup with a Monopolist Landlord

What, then, is holdup? Permit us to explain by examining how opportunism by the landlord and asset-specific investments by coffee bar owners allow holdup to emerge. Recall that under the real estate monopoly, the coffee bar owners expected to cover their long-run average costs because they factored *R* into their calculations and decisions. Believing that they had secure, long-term leases, they decorated their properties and purchased expensive commercial-grade espresso machines, much of whose cost was for installation: a water line needed to be run to the espresso machine and a drain needed to be run from the espresso machine to the waste pipe. Because the decor of each coffee bar and the installation of the espresso machine are sunk costs, and because reinstalling the machine is expensive (it needs to be uninstalled, moved, and then reinstalled), the owner's capital cannot easily be redeployed to some other use. In the parlance of economics, the decor and the espresso machine installation are "specific, sunk investments."

The landlord is now in a position to hold up one or all coffee bar owners. He carefully examines the rental contracts and finds that there is no clause permitting the dedicated plumbing lines. He therefore demands a renegotiation of the contracts in order to appropriate as much revenue as he can from one, or all, of the coffee bars. Because the costs of the decor and the installation of the espresso machine are sunk, the landlord can appropriate up to the difference between each coffee bar's total revenues and its short-run costs, less the salvage value of the espresso machine and the decor—an amount referred to by economists as "appropriable quasi rent." The landlord therefore raises the fixed rent—not *R*—by that amount. Coffee bar owners did not foresee this happening. They will now lose money because they cannot recover the costs of their specific, sunk investments.[41]

A large literature, pioneered by several giants in transaction-cost economics—most particularly Benjamin Klein, Robert Crawford, and Armen Alchian[42] and

[41] The landlord does not need to take all the quasi rent to engage in holdup, but he must appropriate at least part of the difference between long-run and short-run total costs—otherwise it cannot be claimed that holdup is occurring. It should be stressed that a rational agent who engages in holdup will aim to appropriate all of the quasi rent; in the long run the held-up coffee bar cannot cover long-run costs and will exit the market anyway, so why leave money on the table?

[42] Benjamin Klein, Robert G. Crawford & Armen A. Alchian, *Vertical Integration, Appropriable Rents, and the Competitive Contracting Process*, 21 J.L. & ECON. 297 (1978).

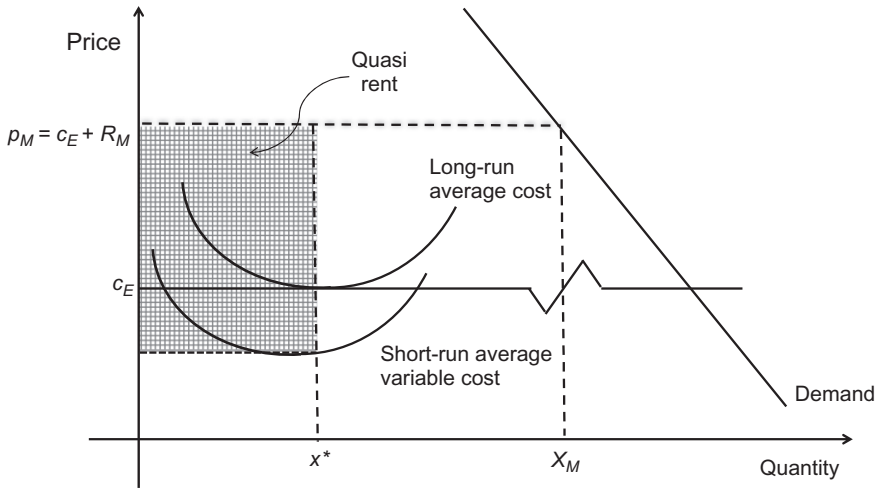

Figure 3. The quasi rent of a coffee shop with a monopoly landlord

Oliver Williamson[43]—pursues the economics of holdup situations such as this one. Their insights are displayed in Figure 3. In that figure we have drawn the standard, *U*-shaped long-run and short-run average cost curves of a coffee bar and the market equilibrium with a monopolist landlord.[44] Like all other coffee bars, this particular coffee bar operates at the point where price, p_M, equals $c_E + R_M$, and it sells x^* espressos (we use lower case letters to denote the sales of individual coffee bars, upper case letters to denote sales in the entire market, and a star to denote that output is at the efficient scale). The landlord knows that when he raises the fixed rent component, the coffee bar will keep selling x^* espressos, at least for a while, because the coffee bar owner cannot easily shift her espresso machine to some other location. The only limit to his opportunism is the coffee bar's short-run costs, because the coffee bar owner will shut down if her monthly rent increases to the point that her total revenues can no longer cover her wage and supply bill. In Figure 3, the quasi rent is indicated by the gridded rectangle.

Note that when the landlord holds up a coffee bar, he now extracts all of the coffee bar's quasi rents (the difference between its total revenues and its short-run costs, which is indicated in Figure 3 with the shaded rectangle). The monopoly rent no longer matters. In fact, as we shall show below, a landlord can hold up the coffee bar owner even if he does not have a real estate monopoly.

[43] Oliver E. Williamson, The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting (Free Press 1985) [hereinafter Williamson, The Economic Institutions of Capitalism].

[44] The *U* shape captures that at very low levels of output the average cost per unit of output is very high, falls as output approaches the optimal scale of production, and then increases again once output has passed that optimal scale. For simplicity, in what follows we ignore the salvage value of the espresso machine.

Note also that the landlord cannot hold up one or all coffee bars over and over again. The coffee bar owners see that they are sheep being fleeced. Once their capital equipment wears out, they will not renew their leases and will exit the market.[45] It follows that holdup and market power are mutually inconsistent economic mechanisms: if the landlord decides to extract all the quasi rents through holdup, he has decided to forfeit exercising market power in the long run—because there will be no market in the long run.

4. *Holdup in a Competitive Real Estate Market*

We cannot stress strongly enough that an input provider does not need market power to engage in holdup. All he needs is another firm that has sunk investments that are specific to him, an incomplete contract, and the willingness to engage in opportunistic surprise. This can be seen in Figure 4, in which the real estate market is once again competitive. A coffee bar in this market operates at the point where its long-run a cost c_E equal its long-run marginal cost, and produces x^* espressos at price p_E. Nothing in this situation prevents any landlord from choosing to raise the fixed rent on the coffee bar in order to extract its quasi rent. As Figure 4 shows, a competitive landlord can hold up coffee bars just as the monopoly landlord can. This basic fact has long been understood in transaction-cost economics. As Benjamin Klein observed many years ago:

> [I]f there is only one potential lessor of office space in a city, then that lessor will be able to charge a monopoly price for office space. If the lease contract is complete or if office-specific investments are not made by lessees, the monopoly lessor will not be able to "hold-up" the lessee at all. If specific investments were made and the lease contract were incomplete, a "hold-up" potential would exist, but this would necessarily be a temporary phenomenon. A monopoly, on the other hand, need not be a short-term phenomenon. . . . Clearly the concepts are distinct.[46]

As we will show, however, this insight has been ignored in the patent-holdup literature.

[45] The point about the non-recurrent, short-run nature of holdup is clearly made by Benjamin Klein:

> This does not mean that "hold-ups" cannot occur. "Hold-ups" do occur from time to time, but only as a short-run phenomenon and only when an unanticipated event leads to a situation where agreed upon contract protection and reputation capital is inadequate. Moreover, when "hold-ups" do occur they will be limited by the fact that buyers will learn to take account of what has occurred and will not make any new reliance investments without taking the necessary precautions, i.e., demanding sufficient reputational or contractual protection.

Benjamin Klein, *Market Power in Antitrust: Economic Analysis After* Kodak, 3 SUP. CT. ECON. REV. 43, 52 (1993).

[46] *Id.* at 59.

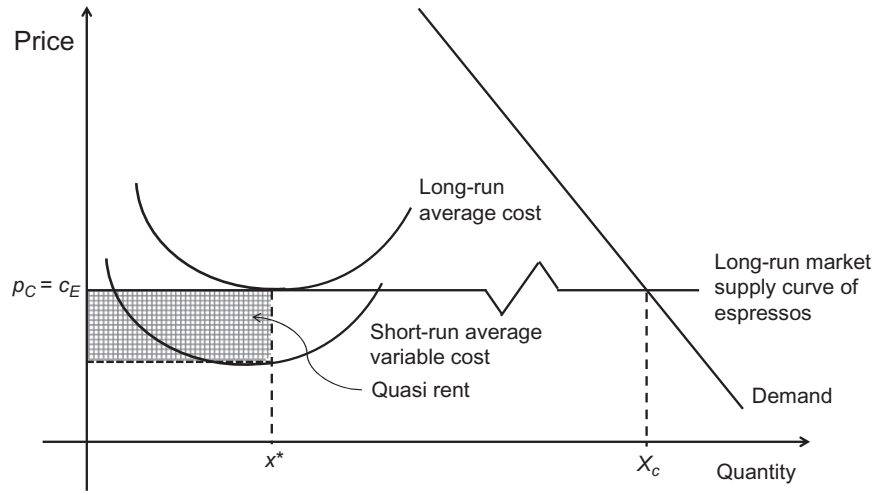

Figure 4. The quasi rent of a coffee shop with a competitive real estate market

B. Why There Is No Holdup Without Opportunism

The transaction-cost literature has always been clear that there are three necessary conditions for holdup to occur. There must be: (1) a sunk investment in a relationship-specific asset; (2) an incomplete contract; and (3) opportunistic surprise.[47] Investment in a sunk, relationship-specific asset is necessary because if the coffee bar owner can easily shift her equipment to another use (for example, by moving it down the street), she can reject the demand for a higher rent. An incomplete contract is necessary because, if every contingency could be contractually anticipated, then there would be no room for renegotiation; any excuse for a rent increase conceived of by the landlord would already be in the contract. Opportunistic surprise—defined by Klein, Crawford, and Alchian as "the unanticipated non-fulfillment of the contract"[48]—is necessary because industries are composed of rational economic agents. The coffee bar owner did not install her espresso machine so that the landlord could appropriate her quasi rents, leaving her with a business that is losing money in the long run.

Klein, Crawford, and Alchian are unambiguous about the need for opportunistic surprise in order for holdup to occur: "After a specific investment is made and such quasi rents are created, the possibility of opportunistic behavior is very real."[49] Klein adds that "hold-ups are always surprises in the sense

[47] For an unambiguous statement of these conditions, see *id.*

[48] Klein, Crawford & Alchian, *supra* note 42, at 300.

[49] *Id.* at 298. There was an acrimonious dispute between Ronald Coase and Klein in the early 2000s about whether holdup requires "deceit," but all transaction-cost theorists agree that holdup requires opportunistic surprise. The dispute about "deceit" arose because of a debate about holdup in the Fisher Body case. *See* R.H. Coase, *The Acquisition of Fisher Body by*

that the particular conditions that will lead to the hold-up are considered unlikely and, because it is costly to negotiate and specify contract terms, these unlikely conditions are not taken into account in the contract."[50]

Williamson, who is widely cited in the patent-holdup literature, also leaves no doubt that opportunistic surprise is a necessary condition:

> How to effect these adaptations poses a serious contracting dilemma, though it bears repeating that, absent the hazards of opportunism, the difficulties would vanish—since then the gaps in long-term, incomplete contracts could be faultlessly filled by recourse to the earlier described general clause device. Given, however, the unenforceability of general clauses and the proclivity of human agents to make false and misleading (self-disbelieved) statements, the following hazards must be confronted: Joined as they are in a condition of bilateral monopoly, both buyer and seller are strategically situated to bargain over the disposition of any incremental gain whenever a proposal to adapt is made by either party.[51]

A reader with an analytic bent might wonder why holdup does not plague virtually every firm in every market. How is it that anything gets produced at all? This was exactly the question that motivated transaction-cost economics. The answer is that firms—both the downstream final producers and the upstream suppliers of necessary inputs—know that holdup is a possibility and that it can, in fact, go both ways (a downstream firm can hold up a supplier if the supplier has made a sunk, relationship-specific investment and there is an incomplete element in their contract). Downstream firms and upstream suppliers therefore devise ways to align one another's incentives such that holdup does not occur, investments continue to be made, and both parties make money in the long run.

Transaction-cost economists therefore studied the numerous adaptations that firms make to avoid holdup. Some adaptations that they focused on are structural; for example, the landlord and the coffee bar owner might vertically integrate by becoming partners. Other adaptations they studied are contractual; for example, the landlord might be required to purchase the espresso machine at its initial cost if he raises the rent beyond some threshold level. Even other adaptions they examined are behavioral; for example, the coffee bar owner might point out that, if the landlord holds her up, she will not rent other space from him in the future, and thus he realizes that it is not in his interest to behave opportunistically. Some of these adaptations might be costly to the firms; others might be so low-cost that the word "adaptation"

General Motors, 43 J.L. & ECON. 15 (2000); Ronald Coase, *The Conduct of Economics: The Example of Fisher Body and General Motors*, 15 J. ECON. & MGMT. STRATEGY 255 (2006); Benjamin Klein, *The Economics Lessons of Fisher Body: General Motors*, 14 INT'L J. ECON. BUS. 1 (2007); Ramon Casadesus-Masanell & Daniel Spulber, *The Fable of Fisher Body*, 43 J.L. & ECON. 67 (2000).

[50] Benjamin Klein, *Why Hold-Ups Occur: The Self-Enforcing Range of Contractual Relationships*, 34 ECON. INQUIRY 444 (1996).

[51] WILLIAMSON, THE ECONOMIC INSTITUTIONS OF CAPITALISM, *supra* note 43, at 63.

is a misnomer—but most adaptations to which they can agree are superior to a situation in which they know that they cannot trust one another, and thus no investment can occur and no money can be made. In short, the mainstream theory of holdup largely explains *successful adaptations* that prevent holdup and sustain trade and investment.

One central policy implication of the mainstream theory of holdup therefore is that the potential problem of holdup should not be an antitrust concern. Williamson suggests that antitrust authorities should not mistake firms' structural and contractual adaptations for the exercise of market power.[52] Klein goes further, concluding that "'hold-up' problems, which are pervasive throughout the economy, do not involve an exercise of monopoly power, and, therefore, are problems for contract law, not antitrust law."[53]

III. FALLACY ONE: PATENT HOLDUP IS A VARIANT OF THE MAINSTREAM THEORY OF HOLDUP

What is the connection between holdup as understood in transaction-cost economics and patent-holdup theory? The short answer is that they contradict one another.

The suggestion that patent holdup is a straightforward variant of the mainstream theory of holdup is not difficult to find in the literature. Shapiro, for example, suggests a connection between patent-holdup theory and the "economics of opportunism," saying that "there is nothing at all exceptional about applying these ideas to patent licensing."[54] In a rebuttal to John Golden,[55] Lemley and Shapiro state:

> An enormous literature explores holdup as a market dysfunction, typically emphasizing the ways in which private firms can manage their affairs to avoid holdup or mitigate its

[52] *Id.* at 16–17 ("At best the administrative apparatus and private ordering supports that attend these transactions [structural and contractual adaptations by firms] are messy. Some scholars decline even to deal with them. Others regard the deviations as evidence of a pervasive condition of 'market failure.' Until very recently the primary economic explanation for nonstandard or unfamiliar business practices was 'monopoly.' . . . That other social scientists should regard these same institutions as antisocial is unsurprising. The enforcement of antitrust from 1945 through 1970 reflected that orientation.").

[53] Klein, *Market Power in Antitrust*, *supra* note 45, at 44. Klein amplifies this view: "Instead of finding 'monopolies' everywhere, hold-up problems should be left to contract law. Antitrust law should not be used to prevent transactors from voluntarily making specific investments and writing contracts by which they knowingly put themselves in a position where they may face a 'holdup' in the future. . . . [C]ontract law inherently recognizes the pervasiveness of transactor-specific investments and generally deals with 'holdup' problems in a subtle way, not by attempting to eliminate every perceived 'holdup' that may arise." *Id.* at 62.

[54] Shapiro, *Patent Reform*, *supra* note 4, at 120.

[55] *See* John M. Golden, *Patent Trolls and Patent Remedies*, 85 TEX. L. REV. 2111 (2007).

> effects. The classic reference in this literature is Oliver Williamson's 1985 book, *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*.[56]

Farrell, Hayes, Shapiro, and Sullivan also suggest a link between patent holdup and holdup as it is understood in transaction-cost economics:

> This article focuses on a problem that the ABA *Handbook* [*on the Antitrust Aspects of Standard Setting*] labels "patent ambush" and that economists call "opportunism" or "hold-up." In very broad terms, opportunism or hold-up arises when a gap between economic commitments and subsequent commercial negotiations enables one party to capture part of the fruits of another's investment, broadly construed.[57]

They quote Williamson: "Emphasizing how parties may inefficiently seek hold-up power, Oliver Williamson famously described opportunism as 'self-interest seeking with guile.'"[58]

The DOJ-FTC joint report also suggests a connection between patent holdup and holdup as it is understood in transaction-cost economics, and it cites Williamson as a reference:

> This type of hold up is a variant of the classical "hold-up problem." The hold-up problem pertains to problems of relationship-specific investment, whereas the hold up contemplated here pertains to standards specific investment. . . . The potential for one party to hold up another party that has sunk investments specific to the relationship may discourage that other party from investing efficiently in the collaboration in the first place.[59]

What are the differences between patent-holdup theory and the established theory of holdup? There are at least two variants of patent-holdup theory, and both elide one of the three necessary conditions of transaction-cost holdup. These elisions allow patent-holdup theory to predict market failure, and call for intervention in markets by government competition authorities. Let us take both variants of patent-holdup theory in turn and explore how they depart from the established theory of holdup.

A. Variant One of Patent-Holdup Theory: Standard-Setting Patent Holdup

In one variant of patent-holdup theory—which, for short, we call "standard-setting patent holdup"—opportunistic surprise is claimed not to be necessary

[56] Mark A. Lemley & Carl Shapiro, *Reply: Patent Holdup and Royalty Stacking*, 85 TEX. L. REV. 2163, 2164 (2007).

[57] Farrell, Hayes, Shapiro & Sullivan, *supra* note 4, at 603–04 (citing ABA SECTION OF ANTITRUST LAW, HANDBOOK ON THE ANTITRUST ASPECTS OF STANDARD-SETTING 60–64 (2004)).

[58] *Id.* at 604 (quoting WILLIAMSON, THE ECONOMIC INSTITUTIONS OF CAPITALISM, *supra* note 43, at 47).

[59] DOJ & FTC, ANTITRUST ENFORCEMENT, *supra* note 9, at 35 n.11. It is interesting to note that Williamson seldom, if ever, uses the term "holdup." For example, it appears neither in the index of *The Economic Institutions of Capitalism*, nor in his Nobel lecture, Oliver E. Williamson, *Transaction Cost Economics: The Natural Progression*, 100 AM. ECON. REV. 673 (2010).

for holdup to occur; all that is needed is a standard-specific investment and an incomplete contract. Consider the formulation provided by Farrell, Hayes, Shapiro, and Sullivan: "We focus on the mechanism of, and techniques for avoiding, inefficient patent hold-up. The pure economics are largely unaffected by whether or not guile is involved, but of course policy and legal treatment may be strongly affected."[60]

The DOJ-FTC joint report appears to follow this line of thinking by ignoring opportunistic surprise in its discussion of patent holdup:

> In the standard-setting context, firms may make sunk investments in developing and implementing a standard that are specific to particular intellectual property. To the extent that these investments are not redeployable using other IP, those developing and using the standard may be held up by the IP holders.[61]

Why is opportunistic surprise left out of the DOJ-FTC formulation? The answer is clear if you parse the sentences. The firms that "may be held up by the IP holders" are freely and openly coordinating around a particular patented technology. They know that other firms own the relevant IP rights, and they know that those patent holders want to earn a royalty for those rights. No one in this situation is likely to be opportunistically surprised—least of all the firms that are "developing and using the standard."

The elision of opportunistic surprise in standard-setting patent holdup matters because, if it is not necessary for one party to opportunistically surprise the other, then holdup could be claimed to be taking place any time that there is a relationship-specific investment and an incomplete contract, which now become sufficient conditions, not merely necessary conditions. Consider the formulation made by Farrell, Hayes, Shapiro, and Sullivan: "Hold-up can arise, in particular, when one party makes investments specific to a relationship before all the terms and conditions of the relationship are agreed."[62] The proof that patent holdup has occurred is simply that a patent holder demands royalties from an unhappy licensee who made a relation-specific investment.

[60] Farrell, Hayes, Shapiro & Sullivan, *supra* note 4, at 604. We are not the first to observe that the definition of holdup provided by Farrell, Hayes, Shapiro, and Sullivan contradicts the established theory of holdup. Epstein, Kieff, and Spulber observe that "'hold-up' is a term of art in the economic literature," and they then note that Farrell, Hayes, Shapiro, and Sullivan and the joint DOJ-FTC report depart from that definition, particularly by eliding guile: "The concept of 'hold-up' has been extensively elaborated on in work by the Nobel Prize-winning economist Oliver Williamson, who also referred to it as 'opportunism,' which he defines a[s] 'self-interest seeking with guile.' The presence of the term 'guile' in this definition is key, and contemplates both that the perpetrator of the behavior acts badly and that the victim is unaware." Epstein, Kieff & Spulber, *supra* note 27, at 15–16 (footnote omitted) (quoting WILLIAMSON, THE ECONOMIC INSTITUTIONS OF CAPITALISM, *supra* note 43, at 61–63).

[61] DOJ & FTC, ANTITRUST ENFORCEMENT, *supra* note 9, at 35 n.11.

[62] Farrell, Hayes, Shapiro & Sullivan, *supra* note 4, at 604.

B. Variant Two of Patent-Holdup Theory: Inadvertent Trespass Patent Holdup

The second variant of patent-holdup theory—which we call "inadvertent trespass patent holdup"—elides the necessity for a contract. Consider the formulation given by Shapiro: "The patent thicket is especially thorny when combined with the risk of holdup, namely the danger that new products will inadvertently infringe on patents issued after these products were designed."[63] The FTC appears to accept this formulation:

> Sometimes, follow-on innovation and production depend[ ] on having access to patents that are economically infeasible to license because they are too numerous to license individually or even to learn about. In other situations, uncertainty surrounding pending patents hampers the reaching of licensing agreements.[64]

Lemley and Shapiro also reframe holdup as inadvertent trespass, while suggesting that the patent owner might behave opportunistically regarding the trespass:

> Consider a downstream firm that is approached by a patent holder who alleges that the downstream firm's product incorporates a feature that infringes its patent. Suppose, for now, that the downstream firm is already selling its product when it learns of the patent claim. This timing may result because the downstream firm designed its product to include a feature for which a patent application was subsequently published or a patent was subsequently issued, perhaps after the patent holder amended its initial claims to capture the downstream firm's product. Alternatively, the downstream firm may simply have been unaware at the time it designed its product that the patent now being asserted had been issued, or it may have been aware of the patent but had no reason to believe the patent owner would argue that the downstream firm's product infringed it. Further, in some cases, the patent holder can engage in strategic delay or concealment, knowing it will be in a stronger bargaining position once the downstream firm has already designed its product incorporating the patented feature.[65]

The problem with the inadvertent-trespass patent-holdup formulation is not only that there is no opportunistic renegotiation of a contract, but also that there is no contract to renegotiate. Bear in mind that in this formulation the problem is not that there are FRAND-committed, declared SEPs (which are covered in the first variant of patent holdup), which would imply a contract, but that there are non-SEPs held by firms that seek to take advantage of the very fact that the existence of the patent is not known to the infringer. The problem with this formulation is that, in order to claim that patent holdup has occurred, all that a researcher or policy maker needs to show is that a manufacturer made an investment that is specific to a patented technology without first securing a license, and that a patent owner then sought damages or an injunction for infringement. This implication is laid bare by Shapiro:

[63] Shapiro, *Navigating the Patent Thicket*, *supra* note 3, at 119.
[64] FTC, To Promote Innovation, *supra* note 8, at 28.
[65] Lemley & Shapiro, *Patent Holdup and Royalty Stacking*, *supra* note 4, at 1995.

"In the model developed here, downstream users who may be infringing a valid patent are subject to hold-up because they must make sunk investments that are specific to using the patented technology."[66] The proof that holdup has occurred is simply that a patent holder demands a royalty that the manufacturer does not want to pay.

C. Both Versions of Patent-Holdup Theory Are Logically Inconsistent

To guide research or policy making, a theory must be logically consistent. The elision of opportunistic surprise in standard-setting patent holdup and the elision of the need for a contract in inadvertent-trespass patent holdup render both variants of patent-holdup theory logically inconsistent. Table 1 compares standard-setting patent holdup with transaction-cost holdup and inadvertent trespass.

Standard-setting patent holdup does not explain why a firm would make an investment specific to a particular standard knowing that it will be held up.[67] Recall that standard-setting patent holdup elides the necessity for opportunistic surprise. One cannot simultaneously maintain that economic agents are rational and that those rational agents invest knowing that patent holders will expropriate their quasi rents.

The resolution of this logical inconsistency is straightforward: the firms making the investments are not being held up at all. Rather, they are willingly and knowingly choosing the technologies that will enter the standard because it is in their long-run self-interest. As we will discuss below, the whole point of innovation—the creation of patented technologies and their commercialization in actual products—is to create or expand a market so that everyone's surplus increases.

Inadvertent-trespass patent holdup does not explain why a firm, knowing that it is likely to trespass on another firm's property right, does not set aside a reserve to cover the risk of that trespass or demand a higher business risk premium. Recall that inadvertent-trespass patent holdup elides the necessity for an incomplete contract, or any contract at all. One cannot simultaneously maintain that patent holdup is a major risk confronted by innovative firms, and that the rational agents who run firms in those same patent-intensive industries are unaware of the risk that they might sometimes infringe on other firms' property rights.

The resolution of this logical inconsistency is also straightforward: the innovative firms are not being held up at all—they are in a long-run

[66] Shapiro, *Injunctions, Hold-up, and Patent Royalties*, *supra* note 4, at 284.

[67] As Spulber explains, the standard-setting process itself generates information about patents that will be included in the standard. Daniel F. Spulber, *Innovation Economics: The Interplay Among Technology Standards, Competitive Conduct, and Economic Performance*, 9 J. COMPETITION L. & ECON. 777, 801 (2013).

Table 1. Transaction-cost holdup and patent holdup compared

	Transaction-cost holdup	**Standard-setting patent holdup**	**Inadvertent trespass**
Sunk and specific investment	Yes, a relation-specific investment	Yes, a standard-specific investment	Yes, investment is specific to the infringed patent
Incomplete contract?	Yes	Yes	No contract
Opportunistic surprise?	Yes (otherwise it is anticipated and parties adapt)	No (manufacturers participated in setting the standard)	Yes (patent holder demands royalties)
Prediction	Structural or contractual adaptation anticipates holdup, prevents it, and sustains trade	Game begins with holdup. No reinvestment or no trade	Game begins with holdup. No reinvestment or no trade

equilibrium. Rather, firms know that they bear the risk of inadvertently trespassing on an intellectual property right in the process of creating a new product, and they either insure themselves against that risk (by setting aside a reserve) or demand a higher expected return on capital, as with any other business risk.

D. A Fundamental Flaw Produced Incomplete Theories

The logical inconsistencies in both variants of patent-holdup theory are rooted in a fundamental flaw in the underlying game theory. As Epstein, Kieff, and Spulber have noted, patent-holdup theory arbitrarily assumes that technology developers and manufacturers negotiate royalty rates only after the technology developers have invested in R&D and the manufacturers have made sunk, standard-specific investments. Until they negotiate royalties, technology developers and manufacturers are in the dark about one another; they decide behind a veil of ignorance.[68]

This view of the process of innovation in SEP-intensive IT products bears little relationship to reality. In practice, technology developers, manufacturers, and other industry participants work together in a protracted, collective manner

[68] Epstein, Kieff & Spulber, *supra* note 27. An exception is the article by Bernhard Ganglmair, Luke M. Froeb & Gregory J. Werden, *Patent Hold-Up and Antitrust: How a Well-Intentioned Rule Could Retard Innovation*, 40 J. INDUS. ECON. 249 (2012). They model a technology developer who invests in R&D, patents her innovation, and then bargains for a license with the manufacturer. If the technology developer and the manufacturer reach an agreement, the manufacturer makes an investment that is specific to the innovation and not contractible. In the last stage of the game the patent holder might hold up the manufacturer. Nevertheless, in equilibrium the innovator invests in R&D, the manufacturer invests in a standard-specific asset, and no holdup occurs. This article is in the transaction-cost holdup tradition, because it models the whole game tree and allows for contracts to solve the holdup problem.

to develop new technologies that support new products that consumers will highly value. They set standards so that all products are compatible across brands and with older versions of those products. The settings in which these collaborations take place are perhaps best understood as technology development organizations, rather than standards development organizations.

This incompleteness of the theory explains why both versions of patent-holdup theory predict market failure. In the established theory of holdup, firms—working together—will make structural, contractual, and behavioral adaptations in order to solve the holdup problem, thereby sustaining trade and investment. In patent-holdup theory, by contrast, firms cannot adapt and solve the problem of opportunistic renegotiation of a contract because the game begins after the R&D is completed and manufacturers invest. Adaptations to prevent holdup have been ruled out by construction.

Both variants of patent-holdup theory therefore make a single, stark prediction: affected industries will stagnate, wither, or die. The policy prescription of patent-holdup theory is that government must intervene to fix the holdup problem.

IV. FALLACY TWO: ROYALTY STACKING MAKES PATENT HOLDUP WORSE

A central claim of patent-holdup theory is that many patent holders might hold up manufacturers simultaneously, a phenomenon that the literature calls royalty stacking. For example, Lemley and Shapiro state that, "As a matter of simple arithmetic, royalty stacking magnifies the problems associated with injunction threats and holdup, and greatly so if many patents read on the same product. In this key sense, the problems of injunction threats and royalty stacking are intertwined."[69] They later amplify this claim: "Not surprisingly, the existence of such 'royalty stacking' exacerbates the holdup problem. Simply as a matter of arithmetic, the problems noted above are greater when the downstream firm faces infringement claims from multiple patent owners."[70]

Farrell, Hayes, Shapiro, and Sullivan state:

> When different parties own many essential patents covering a given standard, the hold-up problem may become more severe, and problems of complementarity arise. The hold-up problem gets worse because collectively the patent holders are apt to take a large proportion of incremental ex post surplus, so investments by technology users are subject to more severe expropriation ex post.[71]

[69] Lemley & Shapiro, *Patent Holdup and Royalty Stacking*, *supra* note 4, at 1993.
[70] *Id.* at 2011.
[71] Farrell, Hayes, Shapiro & Sullivan, *supra* note 4, at 641 (citing Lemley & Shapiro, *Patent Holdup and Royalty Stacking*, *supra* note 4).

Lemley states:

> What are those specific features that facilitate holdup? Insufficient discounting in damages is one. If a patent suit goes to court, the plaintiff may take the patent and, for example, the Intel microprocessor to the jury and say, "You know, they make billions of dollars on this microprocessor. I have a circuit that is used in this microprocessor and all I want is 1%. How can that be unreasonable to ask?" 1% is indeed reasonable in a lot of circumstances. It may not be reasonable, though, if there really are 5000 different inventions bundled together in the microprocessor that Intel sells, because if Intel has to pay 1% 5000 times, it will find it hard to make a profit on its microprocessor.[72]

We show below that there is actually no relationship between holdup and royalty stacking. Indeed, they have different, and to some extent contradictory, empirical implications. The claim that royalty stacking is patent holdup repeated over and over on the same firm is the result of the conflation of holdup and market power.

A. What Is Royalty Stacking?

Royalty stacking is an application of the Cournot complements problem.[73] In 1838, Augustin Cournot asked what happens when two upstream monopolists, one producing zinc and another producing copper, both post their prices to a downstream brass producer independently of one another. He demonstrated that they would post higher prices and sell less than if they would collude and choose a single, profit-maximizing price for both copper and zinc.[74] The implication is that two monopolies are worse than one.

Royalty stacking substitutes patent holders for Cournot's upstream monopolists, and then notes that there might be dozens, scores, or hundreds of such patent holders.[75] Many researchers have called these clusters of patents

[72] Lemley, *Ten Things to Do About Patent Holdup of Standards and One Not To*, *supra* note 4, at 152.

[73] Geradin, Layne-Farrar & Padilla, *supra* note 25, at 145 ("Royalty stacking is at its heart a reincarnation of the 'complements problem' first studied by the French engineer Augustin Cournot in 1838.").

[74] AUGUSTIN COURNOT, RESEARCHES INTO THE MATHEMATICAL PRINCIPLES OF THE THEORY OF WEALTH ch.9 (Irving Fisher trans., MacMillan 1897). Spulber shows that the Cournot complements problem emerges only if input monopolists independently post linear unit prices. (A linear price is a unit price that does not vary with the quantity purchased.) The Cournot complements problem disappears, for example, if manufacturers and input providers bargain bilaterally. Spulber, *Complementary Monopolies and Bargaining*, *supra* note 33. Spulber also shows that royalty stacking emerges only if patent holders post linear royalties and that it disappears if parties bargain for royalties. Daniel F. Spulber, *Patent Licensing and Bargaining with Innovative Complements and Substitutes* 70 RES. IN ECON. 693–713, 2016.

[75] Galetovic, Haber, and Zaretzki note that there are fundamental differences between the market for physical inputs to production, such as copper and zinc, and the market for patented technologies, which makes the application of the Cournot complements framework to patenting highly problematic. Galetovic, Haber & Zaretzki, *Is There an Anti-Commons Tragedy in the Smartphone Industry?*, *supra* note 23.

held by dispersed owners "patent thickets." Some researchers have then assumed that a patent thicket leads inexorably to royalty stacking, and thus to royalty charges to downstream manufacturers that are high relative to the royalties that those manufacturers would have been charged by a single monopolist who owned all patents.[76] Indeed, the patent-holdup literature provides a mathematical expression to operationalize royalty stacking, which Lemley and Shapiro clearly express:

> According to the general theory of Cournot complements, the equilibrium level of output by the downstream firm tends to be smaller the more fragmented the ownership of a given set of patents that read on the downstream product. . . . [I]f marginal costs are constant and the downstream firm faces linear demand, the output level if N essential patents are owned by N separate firms is equal to the output level if all N patents were owned by a single firm multiplied by the factor $\frac{2}{N+1}$. For example, with three patents held by separate firms, downstream output is half as much as it would be if a single company owned all three patents.[77]

We can operationalize the Lemley-Shapiro mathematical expression and plot it in Figure 5. Call N the number of patent holders, $Q^C = 100$ the quantity demanded with marginal cost pricing (the price that would obtain if patent holders did not charge any royalties), and $Q^1 = 50$ the equilibrium quantity when there is a single patent holder. The long-run marginal cost is equal to 10, and the maximum price that a consumer is willing to pay for a device is 110. Assume further that the demand curve is linear, that downstream manufacturers are perfectly competitive, and that each patent holder enjoys market power and sets her royalty individually to maximize profits. Some algebra shows that output is a function of:

$$\frac{1}{N+1}Q^C = \frac{1}{N+1}100.$$

Straightforward substitutions show, for example, that when one patent holder charges a profit-maximizing royalty rate, equilibrium output is cut by half relative to marginal cost pricing, so that $Q^1 = Q^C/2 = 50$. With a second patent holder, the cumulative royalty increases and output falls further to one-third relative to no royalty. By the time the number of patent holders reaches nine, output is 10—which is to say that it is 90 percent lower than with no royalty. And if the number of patent holders is 99, then output would be 1—99 percent lower than with no royalty! Obviously, the equilibrium price rises as output falls, with the increase in price a function of the slope of the demand curve.

[76] According to Egan and Teece, there are at least six different meanings ascribed to the term "patent thicket," only one of which is royalty stacking. Evidence of a thicket, whatever it means, is not necessarily evidence of royalty stacking. Egan & Teece, *supra* note 31.

[77] Lemley & Shapiro, *Patent Holdup and Royalty Stacking*, *supra* note 4, at 2014.

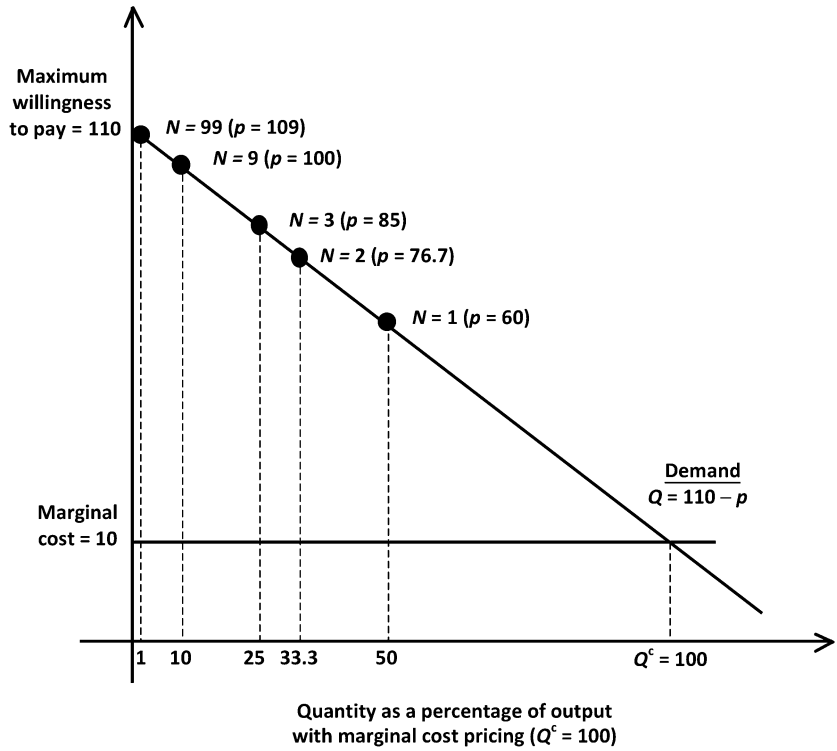

Figure 5. Royalty stacking: equilibrium prices and output with an increasing number of patent holders

The mechanics of royalty stacking mean that it would only take a few patent owners to devastate an industry.[78] High cumulative royalty rates levied on manufacturers mean that they must charge a price for their products that is so high that it will exclude all but a few buyers.[79] Royalty stacking is not, therefore, consistent with a thriving industry: the incentives for incumbent manufacturing firms to invest are weak, the incentives for new manufacturing firms to enter are nil, and the incentives for technology developers are eroded by royalty rates that might not pay for their R&D expenditures.[80] In short, if

[78] The effect is the same with non-linear demand. *See* Galetovic & Gupta, *supra* note 18.

[79] Einer Elhauge explains that before this happens, manufacturers will choose to produce with the old, unpatented technology and royalty stacking will not be observed. The result will be, however, be that technological progress will stagnate. Einer Elhauge, *Do Patent Holdup and Royalty Stacking Lead to Systematically Excessive Royalties?*, 4 J. COMPETITION L. & ECON. 535, 565–66 (2008).

[80] Assume that each patent holder adds a feature that increases the value of the product for consumers. Galetovic and Gupta show that each additional dollar added to the value of the product increases the cumulative royalty by one dollar. In other words, patent holders capture almost all of the additional value. Moreover, as the number of patent holders increases, quantity stagnates and the price that consumers pay increases. Galetovic & Gupta, *supra* note 18.

royalty stacking is actually taking place, then the market will stagnate in the long run.[81]

The empirical implication is straightforward. If a researcher sees a thriving market—one in which there is plenty of investment, new firm entry, and innovation—it either means that there was no royalty stacking in the first place or that market actors mitigated it.

B. Royalty Stacking and Holdup Have Different Empirical Implications

Holdup and royalty stacking have opposite empirical implications when it comes to the prices charged by upstream technology firms to downstream manufacturers. The point of holdup is to extract the appropriable quasi rents of the downstream manufacturers by setting a high royalty for the right to use a patent. By contrast, with royalty stacking, the profit-maximizing royalty rate falls with the number of patent holders. If royalty stacking is actually occurring, the observed cumulative royalty rate is high, but individual royalty rates are low—the opposite of what happens if a patent holder engages in holdup.

This stark result—that individual royalty rates should fall as the number of patent holders rise—can be appreciated with the help of Figure 5. In the figure, the cumulative royalty rate, R, is equal to the equilibrium price (p) minus marginal cost. With one patent holder, $R = 60 - 10 = 50$. With two patent holders, $R = 76.7 - 10 = 66.7$, but now two patent holders split the cumulative royalty, so the individual royalty is equal to $66.7 / 2 = 33.3$. Once there are nine patent holders, $R = 100 - 10 = 90$, the individual royalty is equal to 10. In short, as we move from one patent holder exploiting market power in the setting of a royalty rate to nine patent holders, the individual royalty rate falls by 80 percent!

Some patent-holdup theorists, and even some federal judges, miss this point.[82] But there is nothing controversial about the finding that royalty

[81] It should be stressed that this is the outcome of profit-maximizing behavior by individual patent holders, each setting royalties independently to maximize its individual royalty revenue. The technical reason is explored in detail by Galetovic and Gupta: as the number of patent holders increases, so does the cumulative royalty and the price elasticity of the final demand, and when demand is more elastic, the individual, profit-maximizing royalty is smaller. *Id.*

[82] Contreras and Gilbert, for example, state: "A further consideration in the evaluation of reasonable royalties for both SEPs and non-SEPs arises when, as is often the case, a product requires licenses to use many patents held by different owners. In that circumstance, each patent owner has an incentive to demand a large share of the value of the product and the resulting total royalty demand can exceed the demand that would maximize a licensor's profit if it were the sole source for all of the patents. This phenomenon is called 'royalty-stacking.'" Contreras & Gilbert, *supra* note 7, at 1488. In *Microsoft v. Motorola*, Judge Robart also appears to have gotten the relationship of individual royalty rates and royalty stacking wrong: "Motorola's royalty request for its 802.11 SEP portfolio raises significant stacking concerns. 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

stacking implies low individual royalty rates. Lemley and Shapiro are quite clear about it:

> [T]he aggregate or stacked royalty rate is not simply the sum of the royalty rates that would be negotiated bilaterally by each patent holder in the absence of the other patent holders. . . . The larger are the royalties that the downstream firm is paying to *other* patent holders, the smaller are the margins on the downstream firm's product . . . and the lower is the negotiated royalty rate.[83]

The fact that holdup implies high individual patent royalty yields and that royalty stacking implies low individual patent royalty rates underscores why it is not logically possible to sustain that royalty stacking is patent holdup repeated over and over. If a patent holder is actually holding up all downstream producers in a particular market, then he will obtain a high individual royalty yield and take all the revenue in that market in excess of the downstream firms' aggregate short-run costs. By contrast, if a patent holder is part of a royalty stack, then he will charge a low individual royalty rate. He will share the market power rent—the excess of total revenues over long-run costs—with the other patent holders, but the downstream firms will continue to cover their long-run costs. Recall a fundamental fact that we mentioned in Part II: the total amount of surplus created by economic transactions in a given market is bounded by the demand curve. Royalty stacking and holdup are two mutually inconsistent economic mechanisms; their only commonality is that their goal is to appropriate part of the same economic surplus.

C. Is There a Way Out for Patent-Holdup Theory?

If patent holdup is not actually a variant of holdup as it is understood in transaction-cost economics, and if royalty stacking is not holdup repeated multiple times on the same product, then on what basis can a claim be made that the patent system threatens innovation? One claim that can be made is that, although holdup is not a systemic problem that threatens innovation in SEP-intensive IT industries, there is a potential Cournot complements problem.

A potential Cournot complements problem would not, however, be an antitrust issue. The reason is straightforward: a patent, by design, confers a temporary property right on an inventor—the right to exclude an infringer—and that property right might permit the inventor to exercise a limited amount of market power. Saying that the patent in and of itself generates an antitrust concern would be equivalent to saying that the property right to any

---

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., No. C10-1823JLR, 2013 WL 2111217, at *73 (W.D. Wash. Apr. 25, 2013).

[83] Lemley & Shapiro, *Patent Holdup and Royalty Stacking*, *supra* note 4, at 2011 (emphasis in original).

productive input is sufficient to generate an antitrust concern. In order to claim that antitrust authorities should intervene in a market, one would need to claim something more about the market power of a patent owner. We now turn to the substance of that claim and the fallacy that underpins it.

V. FALLACY THREE: STANDARD-ESSENTIAL PATENTS CONTRIBUTE NOTHING TO THE MARKETS THAT THEY HELP TO CREATE

Patent-holdup theory maintains that the holdup problem is exacerbated when patented technologies are included in the industry standards that make IT products interoperable and compatible. There are actually two variants of this argument. In one variant, it is asserted that when a group of downstream implementers chooses a particular technology as the standard for an industry, it knocks firms that developed alternative technologies out of the market. The result is that, on the supply side, manufacturers are locked into the standard by their own standard-specific investments, whereas on the demand side, consumers would not unilaterally switch to products that use an alternative technology because their devices would no longer be compatible with those owned by other consumers.[84] The firms whose patented technologies have been chosen are now free, at least according to the theory, to charge excessive royalties.[85] This claim—that standard setting confers additional market power and holdup power on patent holders, in excess of the limited market power that necessarily inheres in any patent—opens the door for antitrust and competition authorities to intervene in the process of price setting by the owners of SEPs.

The second variant claims that the value of a standard stems from standardization itself, that the technologies that underpin the standard add little or no value to consumers, and that appropriating the value of standardization is an undue exercise of market power not granted by the patent. Again, the door is open for antitrust and competition authorities to intervene in the process of price setting by the owners of SEPs.

A. Standards, Market Power, and Holdup

By now it should be clear that the claim that the inclusion of patented technologies in industry standards allows SEP holders to charge excessive royalties has nothing to do with holdup as it is understood in the established, mainstream theory. Manufacturers cannot be surprised that they are locked into another firm's technology; they openly and willingly chose that technology, and they

[84] Suzanne Michel, *Bargaining for RAND Royalties in the Shadow of Patent Remedies Law*, 77 ANTITRUST L.J. 889, 891 (2011); Contreras & Gilbert, *supra* note 7, at 1468.

[85] For empirical evidence that inclusion in the standard might not generally lead to market power, see Anne Layne-Farrar & A. Jorge Padilla, *Assessing the Link Between Standards and Patents*, 9 INT'L J. IT STANDARDS & STANDARDIZATION RES. 19 (2011).

did so hoping that it would become wildly popular with consumers. As rational agents, they also understood that when they rejected alternative technologies, they were giving the firm with the winning technology bargaining power in the negotiations over the royalty rate. Why did they do this? They did it because they wanted to expand the market so as to increase their revenues and profits.

Nevertheless, the patent-holdup literature suggests that the additional market power conferred by the inclusion of patented technologies in industry standards is related to holdup. Consider, for example, Shapiro's formulation:

> In essence, any manufacturer seeking to produce a compliant product must obtain a license from each rights holder to avoid facing an infringement action. Inventing around is typically impractical, as it would preclude the manufacturer from claiming that its products are compliant and thus assuring consumers that they are fully compatible with the prevailing standard. Thus, standard setting very often has especially strong elements of both the complements problem [royalty stacking] and the holdup problem.[86]

George Cary, Mark Nelson, Steven Kaiser, and Alex Sistla join holdup, market power, and standard setting in one formulation:

> Selecting a standard ordinarily requires an SSO to choose among competing technologies, and the process frequently results in a collective selection of a patented technology to the exclusion of other patented or non-proprietary technologies. Consequently, standardization necessarily entails the exclusion of alternative technologies . . . . Indeed, because the opportunistic conduct resulting in patent holdup specifically "concerns the inefficient acquisition of market power," many commentators have "generally assumed that [such] opportunism in the standard-setting process is an antitrust problem."[87]

Renata Hesse, speaking as a DOJ official, also links standards, market power, and patent holdup:

> Once a standard becomes established, firms implementing the standard may find switching away more difficult and expensive. This lock-in confers market power on the owners of the incorporated patents. . . . Standards essential patent holders may seek to take advantage of the market power that standardization of their patented technology creates by engaging in hold-up. . . . This type of hold-up raises particular competition concerns when alternative technologies that could have been included in the standard were instead excluded from it.[88]

[86] Shapiro, *Navigating the Patent Thicket*, *supra* note 3, at 128; *see also* Farrell, Hayes, Shapiro & Sullivan, *supra* note 4, at 607; Robert A. Skitol, *Concerted Buying Power: Its Potential for Addressing the Patent Holdup Problem in Standard Setting*, 72 ANTITRUST L.J. 727, 728 (2005).

[87] George S. Cary, Mark W. Nelson, Steven J. Kaiser & Alex R. Sistla, *The Case for Antitrust Law to Police the Patent Holdup Problem in Standard Setting*, 77 ANTITRUST L.J. 913, 914, 921 (2011) (alteration in original) (footnotes omitted) (first quoting Farrell, Hayes, Shapiro & Sullivan, *supra* note 4, at 609; and then quoting Bruce H. Kobayashi & Joshua D. Wright, *Federalism, Substantive Preemption, and Limits on Antitrust: An Application to Patent Holdup*, 5 J. COMPETITION L. & ECON. 469, 471 (2009)).

[88] Hesse, Antitrust and Looking Back on the Last Four Years, *supra* note 11; *see also* DOJ & USPTO, POLICY STATEMENT, *supra* note 11, at 4.

The FTC is even less ambiguous in its conflation of holdup and the potential market power conferred by the adoption of a patented technology into an industry standard: "The ability of patentees to demand and obtain royalty payments based on the switching costs faced by accused infringers, rather than the ex ante value of the patented technology compared with alternatives, is commonly called 'hold-up.'"[89]

B. Do Standards Confer Additional Market Power?

Is there is any substance to the claim that industry standards might confer additional market power on patent holders by knocking down competitors? Once one considers the economics of this claim, it becomes obvious that the notion that industry standards confer additional market power on patent holders requires one to assume that there is little difference between the patented technologies that are adopted into the standard and the patented technologies that are rejected. As we shall show, a fundamental article published in 1962 by Nobel laureate Kenneth Arrow implies that, if this assumption is false, then the edifice of the "industry standard confers additional market power" argument crumbles.[90]

To understand why the theory crumbles if the assumption about similar technologies does not hold, imagine that there is standardized audio-visual technology in use that is not very good, delivering value v_0 per unit. Now imagine that two firms, A and B, have independently created alternative, patented, audio-visual technologies that are better than the old technology, and deliver values v_A and v_B per unit, respectively. Manufacturers consider each patented technology and decide that firm A's technology provides a better entertainment experience for consumers than that of firm B—that is, $v_A \geq v_B > v_0$. In order to make all audio-visual devices compatible with one another, so that the market can expand rapidly, manufacturers adopt firm A's technology as the new industry standard.

Patent-holdup theory maintains that firm A is entitled to at most the incremental difference between the value of its technology and that of the next next-best alternative, which is firm B's technology. The theory therefore postulates that the appropriate price for firm A's technology is what would have emerged had there been a hypothetical *ex ante* price competition between firms A and B.[91] This formulation has been accepted by the FTC:

[89] FTC, THE EVOLVING IP MARKETPLACE, *supra* note 8, at 191.

[90] Kenneth Arrow, *Economic Welfare and the Allocation of Resources for Invention*, *in* THE RATE AND DIRECTION OF INVENTIVE ACTIVITY: ECONOMIC AND SOCIAL FACTORS (Princeton Univ. Press 1962). For an assessment five decades later of the volume in which Arrow made this contribution, see THE RATE AND DIRECTION OF INVENTIVE ACTIVITY REVISITED (Josh Lerner & Scott Stern eds., Univ. of Chicago Press 2012).

[91] *See* George S. Cary, Larry C. Work-Dembowski & Paul S. Hayes, *Antitrust Implications of Abuse of Standard-Setting*, 15 GEO. MASON L. REV. 1241, 1258 (2008); Cary, Nelson, Kaiser

> Courts should recognize that when it can be determined, the incremental value of the patented invention over the next-best alternative establishes the maximum amount that a willing licensee would pay in a hypothetical negotiation. Courts should not award reasonable royalty damages higher than this amount.[92]

Patent-holdup theory also maintains that firm A does not set its price at this appropriate level, $v_A - v_B$. Instead, because manufacturers bilaterally negotiate the royalty rate with firm A after they have collectively chosen it as the standard, they claim that firm A charges the entire difference between the value of the initial technology, v_0, and the value of its technology, v_A—that is, $v_A - v_0$. This royalty exceeds the appropriate value of $v_A - v_B$, because obviously $v_A - v_0 > v_A - v_B$. Manufacturers, therefore, pay higher royalties than they should.

Keep in mind that v_B is a theoretical construct, not a value that is observed in the real world. Indeed, there is no price data for firm B's technology, because that technology was rejected before any products were actually created. Thus, patent-holdup theory claims, on the basis of a hypothetical competition, that whatever royalty firm A charges is excessive. It is not a true statement as an empirical matter; it is not a fact that can even be established. It is true only insofar as it is an implication of the assumptions of the theory.

Now comes the problem: if one takes account of Arrow's insight about the difference between drastic and non-drastic innovations—where a drastic innovation is understood to mean a technology that is much better than the alternative technologies on offer[93]—then it is no longer true as a matter of theory that manufacturers are paying an excessive royalty for the use of firm A's technology. In the example of the audio-visual technologies above, if firm A's technology conferred much more value than firm B's technology—that is, if it pushed out the demand curve for audio-visual products much more than firm B's technology—then economic theory tells us that it will charge less than $v_A - v_B$. Hence, it is not *a priori* true that whatever royalty firm A charges is excessive.

To see why firm A will charge less than $v_A - v_B$ if its innovation is drastic, in Figures 6 and 7 we plot one linear demand curve for firm A's technology, another for firm B's technology, and a third for the old technology.[94] For the moment, it does not matter whether one examines Figure 6 or 7; either will do.

& Sistla, *supra* note 87, at 915; Contreras & Gilbert, *supra* note 7, at 1468; Ramirez, SEPs and Licensing, *supra* note 11.

[92] FTC, THE EVOLVING IP MARKETPLACE, *supra* note 8, at 189.

[93] Arrow studied a process innovation that affected manufacturing cost, but the extension is straightforward. Arrow, *supra* note 90. A brief and clear exposition of Arrow's theory can be found in JEAN TIROLE, THE THEORY OF INDUSTRIAL ORGANIZATION 393 (MIT Press 1988).

[94] Patent-holdup theorists usually assume that demand is perfectly inelastic or that there is a fixed number of units to be licensed, which generates models in which the implications of Arrow's theory and the effect of couching the analysis in terms of demand curves are ignored by assumption. *See* Elhauge, *supra* note 79, at 561–63 (showing that couching the analysis in standard demand theory changes the conclusions of patent-holdup theory).

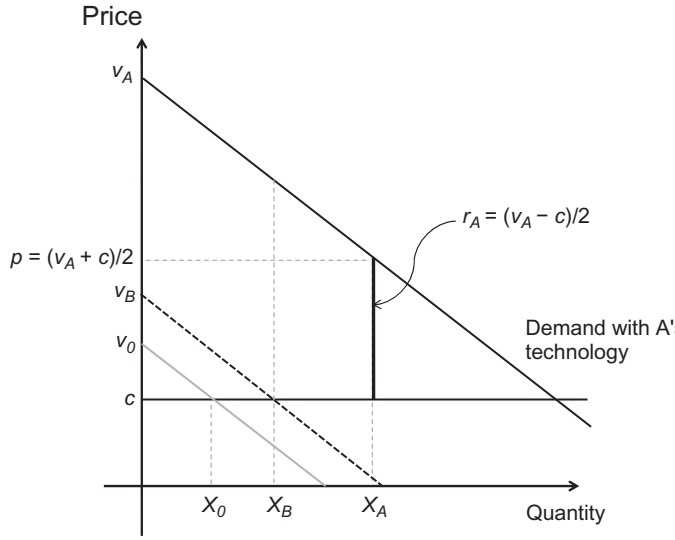

Figure 6. Firm *A* introduces a drastic innovation

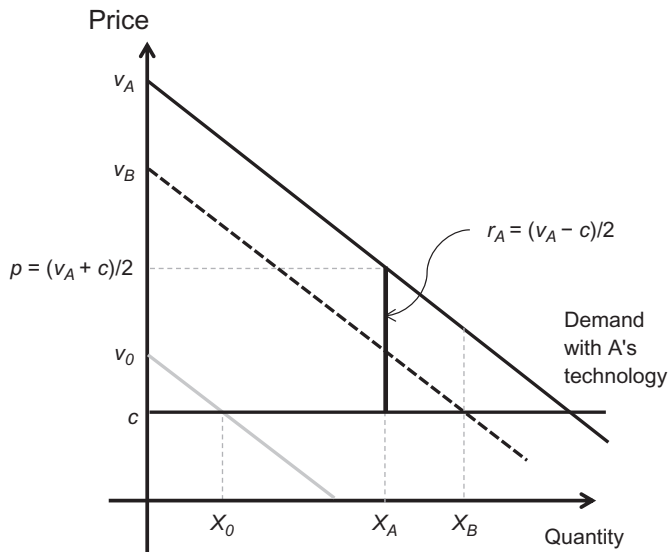

Figure 7. Firm A introduces a non-drastic innovation

We assume that the consumer market is competitive and that the long-run marginal cost of manufacturing audio-visual entertainment devices, c, is constant and the same for all technologies; thus, the supply curve is flat. Because technology A creates more value than does technology B, the demand curve for devices that use technology A (the black continuous line) is higher than

the demand curve for devices that use technology B (the dotted line). Similarly, because technology B creates more value than does the old technology, the demand curve for devices that use technology B is higher than the demand curve for devices that use the old technology (the light-grey continuous curve).

As Figures 6 and 7 show, with the old technology consumers pay c and buy x_0 devices. If the industry adopts firm A's technology as the standard, the demand curve shifts upward because it creates more value than firm B's technology or the old technology, and thus the vertical distance between the demand curves is $v_A - v_0$ and $v_A - v_B$.[95]

What price does firm A set along the demand curve v_A? A little algebra shows that firm A will charge a royalty rate equal to the price-cost margin chosen by a vertically integrated monopoly, the well-known Lerner margin. As indicated in Figures 6 and 7, with linear demand, the profit-maximizing price-cost margin and royalty is

$$r_A = \frac{v_A - c}{2},$$

and thus the equilibrium price increases from c to $p = \frac{v_A + c}{2}$. Note that because consumers value technology A, they buy $x_A > x_0$ units—they purchase many more audio-visual entertainment devices than they did with the old technology. Consumers are better off, which is apparent from the fact that they purchase more units in equilibrium.

Now we get to the central question: how does the royalty rate for firm A's technology (r_A) compare with $v_A - v_B$, the incremental value over its next-best alternative, which is firm B's technology? The answer depends on Arrow's insight: when firm A's innovation is drastic (much more valuable than firm B's) the royalty it charges is actually *smaller* than the incremental value of its technology over firm B's.

To see why firm A charges less than $v_A - v_B$ when its innovation is drastic, examine Figure 6. Note that the vertical distance between the demand curves for devices with technologies A and B is $v_A - v_B$. Note also that demand curve v_A takes a big shift upward compared with v_B because firm A's technology creates much more value for consumers.

What would have happened had there been the hypothetical *ex ante* competition between firms A and B? Imagine that before choosing, manufacturers ask firms A and B to compete on the basis of their royalty rates.

[95] It should be stressed that technology A creates more value than do the alternatives (the existing technology and technology B) because consumers prefer it. This might have to do with the relative engineering merits of the technologies, but it encompasses many other features—for example, design, ease of use, and reliability. This is very important, because the goal of many SDOs is to push out the demand curve as far out as possible by creating products that consumers want.

Because firm B is a rational agent, it knows that its technology is not as good as firm A's, and competition forces it to charge no royalty at all.[96] As Figure 6 shows, if firm B charges no royalty and the industry chooses firm B's technology as the standard, then consumers would pay c for each device and buy x_B units. Firm A's technology provides a better experience for consumers, however, so it can charge $v_A - v_B$, and the industry would sell the same number of devices as when using firm B's technology.

Now things get even more interesting because, as Figure 6 shows, when firm A wins the competition, it wants to charge the royalty rate that would be levied by a profit-maximizing monopolist, $r_A = \frac{v_A - c}{2}$, and that royalty rate is less than $v_A - v_B$.[97] That is to say, firm A would like to charge the same royalty as with no *ex ante* competition. Note that the demand for audio-visual devices shifts upward, consumers buy $x_A > x_B$ devices at a price $\frac{v_A + c}{2}$, manufacturers increase their sales, and firm A earns royalty income. Thus, the claim that whatever royalty firm A is actually charging for its technology is excessive compared with the outcome of a hypothetical *ex ante* competition is wrong.[98]

What if firm A's innovation was non-drastic? Might there then be something to the claim that firm A's royalty rate is excessive? The answer is maybe, but only if one makes strong and unrealistic assumptions about technology development and standard setting.

To see the assumptions necessary to sustain the claim that firm A might be charging more than the "appropriate royalty," examine Figure 7. Note that the difference between firm A's technology and firm B's technology is now small in comparison with Figure 6. The vertical distance between the demand curves is again $v_A - v_B$, but now the demand for devices with firm B's technology is much closer to the demand for devices with firm A's technology. What would be the outcome of the hypothetical *ex ante* competition between firms A and B? Manufacturers would once again choose the best price-quality combination. Because firm A's technology provides a better experience for consumers, it would win the auction charging a shade below $v_A - v_B$, be chosen as the standard, and sell the same number of devices as would firm B. Nevertheless, when the innovation is non-drastic,

$$v_A - v_B < \frac{v_A - c}{2}.$$

[96] *See* Daniel F. Spulber, *How Do Competitive Pressures Affect Incentives to Innovate When There Is a Market for Inventions?*, 121 J. POL. ECON. 1007 (2013) [hereinafter Spulber, *Competitive Pressures*].

[97] A necessary and sufficient condition for $v_A - v_B > r_A = \frac{v_A - c}{2}$ is that $v_A > 2v_B - c$. This defines a drastic innovation.

[98] This is not to take issue with the fact that U.S. courts have used an incremental-value approach to patent valuation for nearly a century. Rather, we take issue with the argument that standardization creates additional market power that enables patent holders to charge more than the incremental value of their innovations.

Therefore, because there is no *ex ante* price competition, Firm A does not charge a shade below $v_A - v_B$; rather, it charges the royalty rate that would be levied by a profit-maximizing monopolist. Consumers pay $\frac{v_A+c}{2}$, which exceeds $v_A - v_B$, and buy only $x_A < x_B$ units.

Now comes the problem: as Epstein, Kieff, and Spulber have observed, this framework requires one to assume that firms A and B do not consider the reward structure of the competition before they invest in all the necessary R&D to develop their competing technologies.[99] They are not allowed to backward induct. As Sidak has made clear, this arbitrarily drawn game assumes that firms A and B are willing to make R&D investments and cooperate in an SDO knowing that, even if they are selected into the standard, they will earn something very close to zero.[100] However, if royalties are routinely driven to amounts close to zero, there cannot be a long-run equilibrium with innovation as an outcome. The reason, as Sidak makes clear, is that innovators expect to make profits that compensate them for the expense and risks of doing R&D in the first place.[101]

The problems with this arbitrarily drawn game are not just theoretical; they are also empirical. For the theoretical construct of a hypothetical, *ex ante* competition between firms A and B to have any bearing at all, there needs to be, as a matter of empirical reality, at least two viable contenders, each of which creates roughly similar value for consumers. Before going any further, a researcher or policymaker who claims that a firm is charging an excessive royalty must show that, at the time a standard was adopted, there were two such technologies in that particular product line, each of which could realistically have created similar value for consumers. If one technology was, as a factual matter, much better than any other feasible alternative, as in Figure 6, standardization could not have created the opportunity to exercise additional market power.

Even if, as a factual matter, there really were two realistic contenders to be the standard in a product line, whether the firm that triumphed in that competition is actually exercising market power is a hypothesis to be tested, not a fact to be assumed *a priori*.

C. The Value of a Standardized Technology and the Value of Standardization

The second variant of the argument that patented technologies add little or no value to the markets they help create distinguishes between the value

[99] Epstein, Kieff & Spulber, *supra* note 27.

[100] J. Gregory Sidak, *The Meaning of FRAND, Part I: Royalties*, 9 J. COMPETITION L. & ECON. 931 (2013) [hereinafter Sidak, *The Meaning of FRAND*]; J. Gregory Sidak, *Tournaments and FRAND Royalties*, 1 CRITERION J. ON INNOVATION 101 (2016).

[101] Sidak, *The Meaning of FRAND*, *supra* note 100; *see also* J. Gregory Sidak, *The Value of a Standard Versus the Value of Standardization*, 68 BAYLOR L. REV. 59 (2016); Sidak, *Tournaments and FRAND Royalties*, *supra* note 100. For an equilibrium model of this process, *see* Spulber, *Competitive Pressures*, *supra* note 96.

created by the standardized technology and the value of standardization itself.[102] It claims that standardization enables interoperability—and thus the development of a network. The technologies that either make the standard work or that use the standard add little, or no, value: any technology, even the existing one, that used this standard would create the same value for consumers. Consequently, any royalty that can be claimed to appropriate the value of standardization is evidence of holdup.

This variant of the theory is similar in many respects to the first variant, and thus is subject to the same criticism. It starts from the same assumptions: firms are not allowed to backward induct; they make costly R&D investments without knowing the reward structure of the competition.

This second variant of the theory, however, is flawed for an additional reason. The claim that patent holders are appropriating the value of standardization requires one to ignore the fact that a standard is an input of production, and inputs are not by themselves sources of utility to consumers. They have value because consumers obtain utility from, and therefore demand, the final product. It only pays to standardize technologies that do things that are valued by consumers; a standard for a useless technology is of no value at all. The value of the standard stems from the demand curve for the final product that is itself valued by consumers.

An everyday, ubiquitous example of a standard illustrates the point. The U.S. dollar is a standard: it is a unit of account and a medium of exchange accepted around the world. It is a very useful input to transacting. Consumers do not, however, obtain utility from knowing that they can avail themselves of a standardized unit of account; they obtain utility from consuming the goods and services they buy with their dollars.

A well-functioning technological standard is like money; it reduces transaction costs, which is valuable, but it does not create value by itself. Maintaining that technology developers should not appropriate the surplus that can appear because consumers can interact seamlessly while using a particular technology is equivalent to saying that sellers should not get any benefit from the reduction of transaction costs wrought by the U.S. dollar.

D. Logical and Empirical Implications

Three implications follow from this analysis. First, any claim that a firm is charging an excessive royalty for a patented technology that is included in an industry standard is a statement about that firm trying to exploit market power; it has nothing to do with holdup, because the crucial element of

[102] For example, in *Ericsson v. D-Link Systems, Inc.*, the Federal Circuit emphasized that a FRAND royalty "must be premised on the value of the patented feature, not any value added by the standard's adoption of the patented technology." Ericsson v. D-Link Systems, Inc., 773 F.3d 1201, 1232 (Fed. Cir. 2014). For a critical analysis, see Sidak, *The Value of a Standard Versus the Value of Standardization*, *supra* note 101.

opportunistic surprise is absent. The downstream firms that are licensing the patents chose a particular technology because they wanted to expand the market, and in doing so they understood that they were potentially improving the negotiating positions of the licensors. Moreover, if a patent holder exploits monopoly power, it will appropriate a monopoly rent and the industry will be in long-run equilibrium; it will not extract a quasi rent, thereby destroying the incentives to invest.

Second, any claim that a firm is trying to exploit market power by charging an excessive royalty for a patented technology that is included in a standard must begin by establishing that the innovation on which its patent reads is non-drastic. Otherwise, the thought experiment of comparing that royalty to the royalty that would be charged on a hypothetical alternative technology is misguided and misleading.

Third, to claim that the inclusion of multiple patented technologies in an industry standard confers market power on multiple patent owners is identical to claiming that there is royalty stacking. Both royalty stacking and the generalized situation of standards conferring market power in any particular industry are identical ways of saying that there are many firms exploiting monopoly power over the same final demand.

The empirical implications are therefore straightforward. If the inclusion of multiple patented technologies that read on an industry standard for a non-drastic innovation is occurring, the results should be the same that one would observe if there was royalty stacking in that industry. Researchers should observe stagnant innovation. That puts us back, however, at the stelae: researchers observe exactly the opposite to be the case. Rates of innovation have run at breakneck speeds, cumulative royalties are low, and firms are induced to enter the manufacture of SEP-intensive IT products. Indeed, if one wanted to identify a set of industries in which innovation was under threat, personal computers, smartphones, and their extension in IoT would be at the bottom of the list of likely candidates.

VI. CONCLUSION

Patent-holdup theory asks an important question: does a decentralized system of technology development—in which the production of complex, interoperable IT products relies on SEPs owned by many firms—stifle innovation, thereby hurting consumers in the form of higher prices and lower-quality products?

There are four possible answers to this question. A decentralized system: (1) promotes innovation and therefore benefits consumers; (2) produces situations that could potentially hurt innovation, but those situations are mitigated by privately ordered solutions at low cost; (3) produces situations that could potentially hurt innovation, forcing firms to make costly, privately ordered adaptations; or (4) produces situations that cannot be mitigated by

privately ordered solutions, thereby harming innovation, or even producing market failure, unless the government intervenes in those markets. Patent-holdup theory, by construction, conjectures the fourth and strongest claim: patent holdup and its associated mechanisms, which are retarding innovation, require government intervention.

Certain fallacies undermine patent-holdup theory: that (1) patent holdup is a straightforward variant of holdup as it is understood in the mainstream field of transaction-cost economics; (2) patent holdup can occur many times over to the same firm, resulting in "royalty stacking"; and (3) patented technologies themselves add little or nothing to the markets that they help create. Underneath those three fallacies is a conflation of two economic concepts, holdup and the exercise of market power. This conflation, and the fallacies that it generates, have produced a system of thought in which the core theoretical claims are assumed to be true *a priori*, rather than considered hypotheses subject to tests against empirical evidence.

The logical flaws of patent-holdup theory, as well as the theory's lack of fit with evidence, leads us to the conclusion that a new theory about the mechanics and dynamics of SEP-intensive IT industries is needed, both as a matter of science and to guide antitrust action and patent policy. Developing that theory is beyond the scope of this article, but any such theory should start from four basic facts that are mostly ignored in the patent-holdup literature: (1) that R&D by technology developers and the setting of industry standards by manufacturers, technology developers, and other stakeholders in SEP-intensive, IT products occurs concurrently and in a protracted fashion; (2) that the development and licensing of technology is characterized by large sunk costs; (3) that technology developers, manufacturers, and other stakeholders play a repeated game in which the technology developers earn reputational rents for being reliable long-run partners of manufacturers and other stakeholders; and (4) that technology developers, manufacturers, and other stakeholders compete with alternative technologies and products. That new theory would explain how repeated play among technology developers, manufacturers, and other stakeholders gives rise to a set of self-enforcing, equilibrium incentives in which output increases, quality improves, prices fall, profit margins attract new entrants and incentivize R&D, patent holders charge royalties well below those predicted by monopoly theory, and consumer welfare increases over time.