

PROGRESS IN THE USEFUL ARTS: FOUNDATIONS OF PATENT LAW IN GROWTH ECONOMICS

Robert D. Cooter* & Uri Y. Hacoheh†

22 YALE J.L. & TECH. 191 (2020)

The U.S. Constitution authorizes Congress to issue patents to promote progress in the useful arts, which we interpret as increasing economic growth through innovation. To ground patent law, we formulate two principles of growth economics. First, selling patents to consume or produce transfers resources to innovating, which speeds growth. Conversely, selling patents to innovate redistributes resources among innovators with deadweight loss, which usually slows growth. Thus, patent protection should be strong against using an innovation to produce or consume, and weak against using an innovation to innovate (separation principle). Second, human welfare can increase exponentially from innovation and quickly overtake any losses from static inefficiency or inequality (overtaking principle). Welfare overtaking is the ethical and political justification of the Constitution's patent clause. Like the Constitution, welfare overtaking suggests that patent interpretation and policy should focus on innovation, not static efficiency or redistribution. Separation and overtaking guide patent law toward its constitutional purpose and increases social welfare from innovation. Alas, in recent years, patent policy has lost its economic foundations. Courts are making doctrinal adjustments that celebrate commercial success instead of innovative superiority (in contrast to the separation principle), and Congress is called upon to reduce static inefficiency concerns and improve consumer access (in contrast to the overtaking principle). This Article criticizes these recent legal developments and proposes several doctrinal adjustments to the law of improvements and experimentation that would bring patent policy closer to its constitutional mandate.

* Herman Selvin Professor of Law, University of California at Berkeley School of Law.

† Fellow in the Center for Law, Economics & Politics (LEAP) at the University of California Berkeley School of Law. We are grateful for comments from Aaron Edlin, Elhanan Helpman, and participants in the seminar of the Center for Technology, Innovation & Competition, University of Pennsylvania Law School; the law and economics seminars at Stanford Law School (including Mitch Polinsky, Mark Lemley, Marus Cole, Michael Klausner, and George Triantis), the University of Virginia Law and Economics Seminar (including Michel Barzuza, Curtis Quinn, John Dufy, and Dotan Oliar); and the law faculty seminar at Florida State University, Washington University in St. Louis, and the University of Arizona Law School.

TABLE OF CONTENTS

I. INTRODUCTION	193
II. ECONOMIC FOUNDATION: GROWTH NOT EFFICIENCY.....	197
A. Business Ventures: Innovation's Engine	197
B. Patents' Effects	199
III. LEGAL FOUNDATION: SEPARATION, FERTILITY, AND OVERTAKING	202
A. Separation Principle.....	202
B. Fertility Principle	205
C. Welfare Overtaking.....	207
IV. OPTIMAL STRENGTH OF PATENTS	214
A. Duration	214
B. Breadth.....	217
C. Remedy	224
V. DOCTRINE AND POLICY.....	228
A. Weakening Patents Against Innovation.....	228
i. Idea Piggybacking	229
ii. Experimenting	230
iii. Improving	236
B. Weakening Patents Against Production.....	244
i. Holdout.....	244
ii. Uncertainty	248
VI. CONCLUSION.....	254

I. INTRODUCTION

Patent policy is losing its economic compass. During the Trump administration, under the banner of curtailing “evergreening,” Congress has seen an unprecedented number of bills suggesting reforming the patent system to prevent innovative companies from scaling up drug prices that burden consumers’ welfare.¹ Trump himself situated the issue at the top of the nation’s agenda in his State of the Union speech in 2018,² then again in 2019.³ Similarly, during the Obama administration, under the banner of curtailing “trolling,” Congress saw an unprecedented number of bills suggesting reforming the patent system to prevent uncommercialized patent rights from imposing a drag on the producers of patented technologies.⁴ Obama himself “became the first president to elevate patent reform to a national meat-and-potatoes issue” when urging Congress to address patent trolls in his 2014 State of the Union speech.⁵ Various academics, reporters, and

¹ See, e.g., Steve Brachmann, *Congress Adds TERM Act and No Combination Drug Patents Act to List of Drug Patent Bills Being Considered*, IPWATCHDOG (June 20, 2019), <https://www.ipwatchdog.com/2019/06/20/congress-term-act-no-combination-drug-patents-act-added-list-drug-patent-bills-considered/id=110525> (mentioning new bills that were proposed during the 116th Congress to deal with evergreening and the rising prices of prescription drugs); Billy Wynne, Josh LaRosa & Alyssa Llamas, *The Drug Pricing Debate Part II: The Many Acts of Congress*, COMMONWEALTH FUND (July 30, 2019), <https://www.commonwealthfund.org/blog/2019/drug-pricing-debate-part-ii-many-acts-congress>.

² *2018 State of the Union Fact Check*, N.Y. TIMES (Jan. 30, 2018), <https://www.ny-times.com/2018/01/30/us/politics/fact-check-sotu.html> (“One of my greatest priorities is to reduce the price of prescription drugs.”).

³ *SOTU Excerpts on Drug Prices: Same Exact Promise as Last Year, Expect the Same Inaction*, PROTECT OUR CARE (Feb. 5, 2019), <https://www.protectourcare.org/sotu-excerpts-on-drug-prices-same-exact-promise-as-last-year-expect-the-same-inaction>.

⁴ See, e.g., Megan Geuss, *Patent Reform Hits Congress Again with Reintroduced Innovation Act*, ARS TECHNICA (Feb. 5, 2015), <https://arstechnica.com/tech-policy/2015/02/patent-reform-hits-congress-again-with-reintroduced-innovation-act>; Matt Levy, *Why the Obama Administration’s Actions Against Patent Trolls Should Make a Difference*, PATENT PROGRESS (Jun. 4, 2013), <https://www.patentprogress.org/2013/06/04/breaking-why-the-obama-administrations-actions-against-patent-trolls-should-make-a-difference>.

⁵ David Kravets, *History Will Remember Obama as the Great Slayer of Patent Trolls*, WIRED (Mar. 20, 2014), <https://www.wired.com/2014/03/obama-legacy-patent-trolls>.

other stakeholders have followed suit, describing the patent system as a “mess,”⁶ “broken,”⁷ “rotten,”⁸ and other similar superlatives.⁹

Although issues such as the escalating prices of prescription medicine or abusive litigation by nonpracticing entities (NPEs) are of outmost social importance, they are also ancillary to patent theory, policy, and reform. The United States Constitution authorizes Congress to make patent law to “promote the progress of science and the useful arts.”¹⁰ Only by causing *progress* does patent law fulfill its constitutional purpose. To measure progress, economics uses cost-benefit analysis, net national product, the quality of life, and similar indexes. To predict progress, economics uses growth theory. Thus, growth economics explains the extent to which intellectual property law fulfills its constitutional purpose in the United States and some other countries.¹¹

⁶ See *A Question of Utility*, ECONOMIST (Aug. 8, 2015), <http://www.economist.com/node/21660559>.

⁷ See ADAM B. JAFFE & JOSH LERNER, INNOVATION AND ITS DISCONTENTS: HOW OUR BROKEN PATENT SYSTEM IS ENDANGERING INNOVATION AND PROGRESS, AND WHAT TO DO ABOUT IT (Princeton Univ. Press 2004).

⁸ *Time to Fix Patents*, ECONOMIST (Aug. 8, 2015), <https://www.economist.com/leaders/2015/08/08/time-to-fix-patents>.

⁹ See, e.g., Richard A. Posner, *Why There Are Too Many Patents in America*, ATLANTIC (July 12, 2012), <http://www.theatlantic.com/business/archive/2012/07/why-there-are-too-many-pa-tents-in-america/259725> (arguing that the system has “serious problems.”); DAN L. BURK & MARK A. LEMLEY, THE PATENT CRISIS AND HOW THE COURTS CAN SOLVE IT 167–70 (2009) (noting that the patent system is in a “crisis”); William Hubbard, *The Competitive Advantage of Weak Patents*, 54 B.C. L. REV. 1909, 1913 (2013) (“[P]atent protection in the United States should be weakened.”); M.G. Siegler, *Google on the Nortel Loss, Patents as Government-Granted Monopolies, and Plates of Spaghetti*, TECHCRUNCH (July 25, 2011), <http://techcrunch.com/2011/07/25/google-patent-fight> (“A patent isn’t innovation. It’s the right to block someone else from innovating.”); MICHELE BOLDRIN & DAVID K. LEVINE, AGAINST INTELLECTUAL MONOPOLY (1st ed., Cambridge Univ. Press 2008); *Patents: Last Week Tonight with John Oliver*, (HBO Apr. 19, 2015), available at https://www.youtube.com/watch?v=3bxcc3SM_KA; David Lee Johnson, *Facing Down the Trolls: States Stumble on the Bridge to Patent-Assertion Regulation*, 71 WASH. & LEE L. REV. 2023 (2014).

¹⁰ U.S. CONST. art. I, § 8, cl. 8.; *Mazer v. Stein*, 347 U.S. 201, 219 (1954) (“The economic philosophy behind the clause empowering Congress to grant patents and copyrights is the conviction that encouragement of individual effort by personal gain is the best way to advance public welfare through the talents of authors and inventors in ‘Science and useful Arts.’”); *Graham v. John Deere Co.*, 383 U.S. 1, 5–6 (1966) (similar); Alan Devlin & Neel Sukhatme, *Self-Realizing Inventions and the Utilitarian Foundation of Patent Law*, 51 WM. & MARY L. REV. 897, 901 (2009); David S. Olson, *Taking the Utilitarian Basis for Patent Law Seriously: The Case for Restricting Patentable Subject Matter*, 82 TEMP. L. REV. 181, 181 (2009).

¹¹ Some constitutions related intellectual property rights explicitly to “economic development” or “growth” (e.g., Angola, Brazil). Some constitutions also mention

The foundations of patent law in growth economics are easily explained but not widely understood. In this Article, we fill this glaring gap. Patents are served to increase an innovation's price, which transfers wealth from buyers to the innovator. The transfer's effect on growth depends on whether the buyer uses the innovation to consume, produce, or innovate. Consuming and producing are static activities, whereas innovation is a dynamic activity. If buyers use the innovation to consume or produce, the patent transfers wealth from the buyers' static activities to the seller's dynamic activity. The transfer increases the average profitability of innovating, which causes more innovating. Thus, to promote growth, patent protection should be strong against using an innovation for consuming or producing. Unlike the simplified patent troll debate, for example, in the absence of aggravating circumstances, NPEs' enforcement of uncommercialized patents against the producers of patented technology generally promote growth.¹²

Conversely, the effect is different when the buyer uses the innovation to innovate. In that case, buyer and seller are both innovators. The patent redistributes wealth from the buyer's dynamic activity to the seller's dynamic activity. Some resources are lost through redistribution ("deadweight loss"), and the average profitability of innovating usually decreases.¹³ The decrease in profitability causes less investment in innovating and slower growth. Thus, to promote growth, patent protection should be weak against using an innovation for innovating. For example, all other things equal, NPEs generally degrade growth when they enforce patents against innovators, experimenters, or improvers of patented technology.

In sum, the first principle of the economics of patents is the *separation principle*: to promote progress in the useful arts, patent protection should be strong against using an innovation to consume or produce, and weak against using an innovation to innovate. This Article uses the separation principle for prediction, evaluation, and interpretation of contemporary patent policy. The separation principle predicts that strong patents against consuming or

purposes beyond economics, such as the "interests of society" (Angola) or "social interests" (Brazil). Also, some constitutions create IP rights without mentioning their purpose (Bhutan, Burundi, Canada). See generally Lior Zemer, *The Constitutional Fallacy of Intellectual Property Clauses*, 35 BERKELEY J. L. & TECH. (forthcoming 2020).

¹² NPEs may nevertheless degrade growth when they aggravate a holdout problem or uncertain rights. See *infra* Section V.B.

¹³ In economics, "deadweight" refers to a loss without an offsetting gain. The legal interpretation of the constitutional purpose requires focusing on innovation, not other values.

producing promote growth, and strong patents against innovating retard growth. Meaningful studies of the effect of patents on innovation must distinguish these activities. Unfortunately, many legal discussions of the effects of patents on growth are unilluminating because they combine the positive effect of patents against consuming and producing with the negative effect of patents against innovating. Such studies reveal little about the contribution of patents to the rate of innovation.

In addition, conventional legal discussions unnecessarily complicate the evaluation of the welfare effects of patents. A patent enables its owner to increase the price of the innovation above the cost of production. Pricing above cost is the static inefficiency caused by monopoly. However, pricing above cost yields profits that are an incentive to innovate. This is the tradeoff between access to the product and incentives to innovate. Thus, the conventional normative analysis of patents balances access and incentives. A patent increases human welfare when the gain from faster innovation exceeds the loss from static inefficiency.

Instead of this balancing test, a simpler justification usually suffices. As we will show, patent law often applies to conditions where the welfare gains from faster innovation quickly overtake the welfare losses from static inefficiency. We call this proposition the *overtaking principle*. In these circumstances, balancing innovation and static efficiency is unnecessary. We can evaluate patent law by focusing on the gains from innovation and ignoring the losses from static inefficiency. Thus, in contrast to the simplified debate about “evergreening,” in the absence of aggravating circumstances, escalated consumer products’ prices should not justify a patent reform.¹⁴

Turning from evaluation to interpretation, progress is patent law’s constitutional purpose, which we interpret as growth in welfare from innovation. Given two possible interpretations of a patent law, the interpretation that causes more innovation satisfies its constitutional purpose more fully. The interpretation that satisfies a law’s constitutional purpose most fully is often its correct interpretation. Thus, by identifying the growth-maximizing interpretation of a patent law, the separation principle helps to find the correct interpretation.

This Article identifies, coins, and develops the separation principle of patent law and the overtaking principle of welfare analysis. In Part II, we contrast growth and efficiency as alternative

¹⁴ But see Uri Y. Hacothen, *Evergreening at Risk*, 34 HARV. J. L. & TECH. (forthcoming 2020) (justifying a patent reform based on the aggravated circumstances created by complex regulation and health market deficiencies).

foundations of patent economics. In Part III, we use growth economics to explain the separation and overtaking principles. In Part IV, we apply these principles to the three dimensions of patent strength: duration, breadth, and remedy. Finally, in Part V, we show how separation is translated into legal policy and recommend further doctrinal adjustments. We also mention exceptions to the separation principle, such as in cases of holdout and uncertain rights, but emphasize that these cases should remain secondary considerations. A short conclusion follows.

II. ECONOMIC FOUNDATION: GROWTH NOT EFFICIENCY

To develop patent economics, two different economic traditions are available. The first is efficiency economics. Economies become more efficient by allocating resources better, which improves the mix of the same goods made using the same techniques. Because the goods and techniques are the same, efficient resource allocation is called “static efficiency.” Efficiency economics is already assimilated into legal scholarship, and it is mathematically elegant. The second tradition is growth economics. Economies grow by innovating, which makes better goods using better techniques. Compared to efficiency economics, dynamic economics is unassimilated into legal scholarship, and it is mathematically inelegant. However, for patent law, growth economics is far more important than efficiency economics.

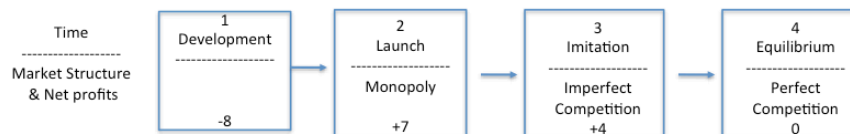
A. *Business Ventures: Innovation’s Engine*

In efficiency economics, the basic analysis focuses on the competitive firm. In growth economics, however, the basic analysis focuses on the business venture, as we will explain. A bold ship’s captain in seventeenth-century England proposes that investors outfit a ship for a voyage to Asia for spices.¹⁵ The voyage is costly and risky, but if it succeeds, the spices will be worth a fortune. Seventeenth-century spice voyages involved up-front investment, high risk, and high return. Similarly, an engineer in Silicon Valley in 2010 has an idea for a new computer chip. Development is costly and risky, but if it succeeds, it will be worth a fortune. Like seventeenth-century spice voyages, twenty-first-century technological innovations involve up-front investment, high risk, and high return.

¹⁵ This characterization of the spice trade is based on Ron Harris, *Law, Finance and the First Corporations*, in GLOBAL PERSPECTIVES ON THE RULE OF LAW 145 (James J. Heckman, Robert L. Nelson & Lee Cabatingan eds., 2009).

A profitable business venture often has a life cycle like the one depicted in Figure 1. The venture begins with the development of a new idea in period 1, which costs 8. By developing the idea into a product, the innovator acquires a valuable secret or patent, or perhaps a cluster of secrets and a portfolio of patents. Because patent protection requires disclosing the invention to the public, the innovator must choose between secrecy and patent protection.¹⁶ The innovator makes this choice after developing the innovation in period 1. After development, the innovation is launched in period 2, which involves marketing the innovation to buyers. When launched in period 2, the innovation has no competitors, so the innovator is a monopolist who receives a payoff of 7. The extraordinary payoff from secrecy or a patent is called “rent” by economists. In period 3, imitators develop substitutes for the innovation. Substitutes provide competition in period 3 that reduces the venture’s profits to 4. Economists say, “The rent from secrecy or a patent dissipates.” If competition proceeds to its logical extreme, the substitutes eventually become perfect, and the market becomes perfectly competitive. With perfect competition, the innovator’s payoff is the normal rate of return, which economists call “zero profits,”¹⁷ as depicted in period 4.

Figure 1. Life Cycle of a Profitable Venture



Summing over the life cycle in Figure 1, the venture’s overall profitability equals +3. Thus, Figure 1 depicts the life cycle of a profitable venture. A venture’s profitability is a fraction of its social value—usually a small fraction. The rest goes to consumers

¹⁶ Patents are an alternative to trade secrets because they require disclosure. See Note, *Patent Preemption of Trade Secret Protection of Inventions Meeting Judicial Standards of Patentability*, 87 HARV. L. REV. 807, 811 (1974) (“This exchange of secrecy for exclusivity is a principal foundation of the patent system.”); Gordon L. Doerfer, *The Limits on Trade Secret Law Imposed by Federal Patent and Antitrust Supremacy*, 80 HARV. L. REV. 1432, 1441 (1967).

¹⁷ Perfect competition drives the prices of all goods to their cost of production. Profits are zero after including the cost of capital in the other costs of production. The cost of capital equals the ordinary rate of profit in alternative uses.

as “surplus”¹⁸ and to other firms as profits. Thus, the wealth that Apple investors obtained from the iPhone is a fraction of its value to consumers, imitators, and applicators.¹⁹

In period 4 in Figure 1, production continues and competition drives profits to zero. At the end of period 4, production ceases because the product is obsolete. A new innovation destroys the old one’s value. When ventures like the one illustrated in Figure 1 repeat themselves, innovation grows on itself, and social welfare spirals up like the falcon’s gyre.

B. Patents’ Effects

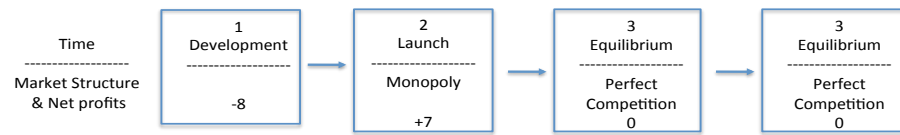
Figure 1 depicts a profitable venture. Before profits fall to zero in period 4, investments are recouped, and more. However, most ventures fail and lose money. Recent U.S. data suggests that 40% of new businesses survive and 60% disappear within four years.²⁰ Figure 2 depicts a losing venture. The innovator in Figure 2 spends 8 in period 1 to develop the product. When the innovation is launched in period 2, the innovator has no competitors and enjoys profits of 7. The only difference between Figures 1 and 2 is in period 3. In period 3 of Figure 1, the innovation yields a payoff of 4 because it is better than its imitations. In period 3 of Figure 2, however, the innovation yields a payoff of 0 because it is no better than its imitations. Thus, the market reaches a competitive equilibrium in period 4 of Figure 1 and in period 3 of Figure 2. Summing over the life cycle in Figure 2, the venture’s net payoff equals -1. Thus, the innovator in Figure 2 cannot recover the cost of developing the new idea. The innovator who foresees the path in Figure 2 will not develop the innovation.

¹⁸ By definition, the consumer’s surplus equals the difference between the price that a consumer would be willing to pay for a good and the price that he or she actually pays.

¹⁹ See generally Zvi Griliches, *R&D and Productivity: Econometric Results and Measurement Issues*, in THE HANDBOOK OF THE ECONOMICS OF INNOVATION AND TECHNOLOGICAL CHANGE (1995).

²⁰ Forty-four percent of U.S. businesses started in the 1990s still existed four years afterward. See Amy E. Knaup, *Survival and Longevity in the Business Employment Dynamics Data*, MONTHLY LAB. REV. 50, 52 (May 2005).

Figure 2. Life Cycle of an Unprofitable Venture



The innovator has the advantage in Figure 1 and the disadvantage in Figure 2. The difference is ease of imitation. The first firm to develop and market an innovation often enjoys market power temporarily. The duration of market power depends on the speed at which competitors imitate. Delaying imitation can transform a losing venture like the one illustrated Figure 2 into a winning venture like that depicted in Figure 1.

Precedence conveys advantage in three ways that affect the speed of imitation. First, secrecy often slows the invention's dissemination to competitors.²¹ Second, the innovator might get big first and enjoy increasing returns to scale. With increasing returns to scale, first in time is first in market might.²² Third, patents create market power by legal fiat. With patents, first in time is first in property right.

Law affects the profitability of all phases of a venture—finance, development, marketing, and competition. The three sources of market power for innovators (secrecy, scale economies, and legal fiat) relate especially to three bodies of law: trade secrets, antitrust, and patent law. With respect to secrecy, containing information is a fundamental purpose of a corporation in a dynamic economy. Trade secrecy laws and corporate incentives help the firm to slow dissemination of its profitable discoveries.

²¹ Trade secrecy works better with explicit information that is easy to copy, like a recipe or a computer code, than with implicit information that is irreducible to simple communication. See generally Yuval Feldman, *Confidential Know-How Sharing and Trade-Secrets Laws: Studying the Interaction between Legality, Social Norms and Justice among High-Tech Employees in Silicon Valley* (2004) (unpublished Ph.D. dissertation, University of California at Berkeley) (on file with author).

²² Market capture may be even more dominate in an environment where there are strong standardization effects in the market, so the standard, once adopted, is likely to be durable. IP rights reinforce market power in such circumstances. See, e.g., Mark A. Lemley, *Antitrust and the Internet Standardization Problem*, 28 CONN. L. REV. 1041 (1996); Peter S. Menell, *Economic Analysis of Network Effects and Intellectual Property*, in 1 RESEARCH HANDBOOK ON THE ECONOMICS OF INTELLECTUAL PROPERTY LAW 156 (2019).

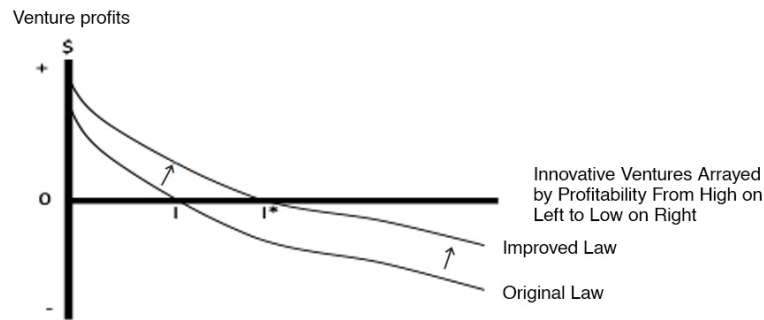
With respect to antitrust, the profitability of innovation for a firm increases by allowing it to monopolize a market for the innovation, which increases its profitability. Also, free trade widens markets, which spreads the development costs of each innovation over more buyers.²³ With respect to intellectual property, patent law (when effective) creates temporary monopolies, as we will discuss at length later.

When better laws make more ventures profitable, more innovations are developed. Figure 3 depicts this fact. Imagine an array of new ideas that differ according to the expected profitability of developing them.²⁴ The vertical axis indicates the profitability of each venture. The horizontal axis arrays ventures by profitability from high on the left to low on the right. Positive profit expectations launch innovative ventures, and negative profit expectations stifle ventures. Given open competition, investors will finance ventures until profits fall to zero. Profits fall to zero where the curves in Figure 3 intersect the horizontal axis. The two curves in Figure 3 contrast venture profits under two legal regimes. Under the original law, venture profits reach 0 at innovation I. Improved law changes the situation by shifting the venture profit curve up, as indicated by the arrows, so venture profits reach 0 at innovation I*. Thus, the improved law increases the number of innovations that get developed from I to I*. Figure 3 depicts the increase in innovations in one period of time. If better law increases innovations in every future period of time, then the growth rate increases.

²³ The size of markets determines the extent of the division of labor, as famously observed by Adam Smith, and a finer division of labor is more efficient. To this proposition about static efficiency, we add that size of markets determines the profitability of innovation. The larger the number of sales for each innovation, the more development costs can be spread over the user base. Thus, volume drives innovation, or free trade drives growth. Laws that affect the size of markets include free trade, global finance, advertising and marketing regulations, and infrastructure development. Also, Spulber explains that antitrust law generally permits vertical integration, and vertical integration helps firms to appropriate the benefits of innovations that patents do not protect fully. 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).

²⁴ Strictly speaking, venture profits equal the discounted present value of the stream of revenues from the innovation's sales, minus the discounted present value of its development and production costs.

Figure 3. More Venture Profits, More Innovations



III. LEGAL FOUNDATION: SEPARATION, FERTILITY, AND OVERTAKING

Ventures are the economic foundation of innovation. Now we discuss the principles that are the legal foundation for patent law.

A. Separation Principle

By blocking competitors, a patent enables the seller to raise the price, which transfers wealth from buyers to the seller. The original seller is the inventor who received the patent. Buyers use the innovation to consume, produce, or innovate. Thus, the patent transfers wealth to innovating that comes from consuming, producing, or innovating.

The transfer of wealth can increase or decrease innovation, depending on how the buyer uses it. When market power transfers wealth from consuming and producing (static activities) to innovating (dynamic activity), the total profits of innovating increase. The increase in incentives to innovate causes faster growth. Thus, to promote progress in the useful arts, the law should give innovating strong property rights against consuming and producing.

Consuming and producing, however, are not the only uses of innovations. In addition, innovations are used for innovating. Many innovations need prior innovations like a carriage needs wheels. When innovators buy and sell to each other, Professor Lemley observes, “It is not enough to say that intellectual property law

favors ‘creators’—for here we have creators on both sides of the equation, and the law must choose between them.’²⁵

How can this choice be made? To promote progress in the useful arts, the law should give innovators weak property rights, or none, against innovating by others. Here is why. Part of the buyers’ losses transfer to the seller as higher profits, and part gets lost without transferring to anyone. “Deadweight” describes a loss without an offsetting gain. Thus, if a monopoly causes, say, the seller to gain 100 and buyers to lose 130, then the difference of 30 is the deadweight loss.²⁶ Given deadweight loss, innovators gain less on average from selling patents to each other than they lose from buying them. Deadweight loss decreases the average profitability of innovating, which slows growth.

To illustrate, consider a licensing fee to use pharmaceutical drugs in research. For testing and developing, newer drugs often use older drugs. Patents increase revenues from innovators selling drugs to other innovators, and patents decrease revenues from innovators buying drugs from other innovators. In a chain of transactions, the latter exceeds the former. The decrease in total profits of innovators is the deadweight loss among them. When total profits of innovating decrease, so do investments in innovating. Thus, a patent right for pharmaceutical drugs against use by others for testing and developing new drugs should slow the rate of innovation.

In sum, for promoting innovation, the ideal strength of a patent depends on the activity burdened with higher prices. The *separation principle* asserts that *the innovator’s rights should be strong against others consuming the innovation or producing with it—broad and long with a generous remedy for infringement. Conversely, the innovator’s rights should be weak against others using the innovation to innovate—narrow and short with little remedy for infringement.*²⁷

²⁵ Mark A. Lemley, *The Economics of Improvement in Intellectual Property Law*, 75 TEX. L. REV. 989, 998 (1997).

²⁶ Note that in the standard graphical analysis, the transfer is a rectangle and the deadweight burden is a triangle. The standard analysis is somewhat more complicated than these remarks suggest because of the difference between long-run and short-run effects. In the long run, competition tends to be perfect, prices tend to equal costs, and all producers receive zero profits. Consequently, all of the long-run cost increase must fall on consumers, not producers.

²⁷ In a seminal work, Judge Pierre N. Leval explained that copyright protection should be weakened against uses that are transformative. Leval has recognized that the separation function should be at the core of the copyright infringement analysis. Pierre N. Leval, *Toward a Fair Use Standard*, 103 HARV. L. REV. 1105, 1111 (1990). According to Leval, to “fulfill the objective of copyright law [and] to stimulate creativity for public illumination,” the question of infringement “turns

Figure 4 depicts the separation principle. Innovator A creates innovation α and sells it to innovator B. B uses it to create innovation β and sells it to consumer C. If B patents β and raises its price, then innovator B extracts wealth from consumer C, which provides an incentive for more innovation. According to the separation principle, to increase innovation, law and policy should enhance the market power of innovators like B against consumers like C.

Figure 4. Three-Party Straight-Line Transaction

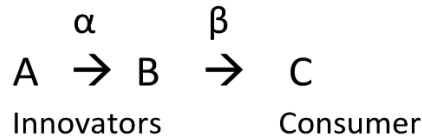
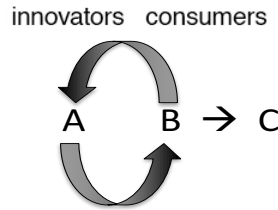


Figure 4 depicts a three-party straight-line transaction from A to B to C. Alternatively, each one may repeatedly sell innovations to the other. To revise the figure accordingly, assume that A sometimes sells innovations to B, and B sometimes sells innovations to A. These, transactions are represented as circular in Figure 5. With circular transactions, the increase in prices from patents cause A and B to gain as sellers and to lose as buyers. They usually lose more than they gain (deadweight loss), so their patents against each other lower the total profitability of innovating. Lower profits among innovators lower incentives to research and develop innovations. Consequently, according to the separation principle, law and policy should not enhance innovator A's market power against innovator B, or innovator B's market power against innovator A.

primarily on whether, and to what extent, the challenged use is transformative.” “[S]econdary use [that] adds value to the original . . . transformed in the creation of new information, new aesthetics, new insights and understandings” should not infringe copyright because “this is the very type of activity that the fair use doctrine intends to protect for the enrichment of society.” *Id.* Leval’s insight is equally important to patent policy. See also Pamela Samuelson, *Freedom to Tinker*, 17 THEORETICAL INQUIRIES IN L. 563, 600 (2016) (“IP rules should be interpreted, or as necessary adapted, so that they are applied flexibly in a manner that promotes the ongoing progress of science and useful arts and other legitimate interests of those who expand horizons of knowledge through tinkering and sharing the results of their tinkering with others.”).

Figure 5. Circular Transactions among Innovators



When marketing to each other, innovators sometimes apply the separation principle. For example, Sun Microsystems developed the Java programming language in the 1990s and made it available for free to developers under a general-purpose license.²⁸ Developers use Java to write programs that they sell to consumers.²⁹ Thus, Java is free for innovators and costly for consumers of its applications. Legal application of the separation principle would cause courts to mimic these licensing practices by private firms.

The preceding discussion assumes that buyers use a patent for consuming, producing, or innovating. In reality, however, many buyers mix uses of the same product. Patent law should separate them and treat them differently. Patent protection should be strong against the same buyer using an innovation for consuming or producing, and weak against the same buyer using the innovation for innovating. When separating them, ordinary language and economic analysis both provide a basis for making the required distinctions. Later, we discuss how some patent laws separate these uses.

B. Fertility Principle

We have explained that redistribution among innovators usually reduces the total amount of innovation. Sometimes, however, redistribution among innovators has the opposite effect: it increases the amount of innovation. To explain why, we distinguish

²⁸ *Sun to GPL Java*, THEOPENFORCE.COM (Nov. 13, 2006, 3:59 PM), http://www.theopenforce.com/2006/11/sun_to_gpl_java.html (“Sun will follow a dual license policy . . . enabling Java to be licensed under either GPL 2 or Sun’s own commercial license which provides legal indemnification.”). For a comprehensive review of Sun’s business strategy, see Peter S. Menell, *Rise of the API Copyright Dead: An Updated Epitaph for Copyright Protection of Network and Functional Features of Computer Software*, 31 HARV. J. L. & TECH. 305, 347–55 (2018). A reverse of this pattern sometimes occurs as when the owner of a platform for running computer games allows consumers to use it for free and charges developers who want to use it to create new games to sell to consumers.

²⁹ Legal disputes subsequently arose when Microsoft modified Java to run exclusively on its Windows program, so only Windows users could run programs in modified Java. See Menell, *supra* note 28, at 353.

among innovations by fertility. Today's new molecule is discovered from yesterday's new molecule; today's new operating system is discovered from yesterday's new operating system; today's new power cell is discovered from yesterday's new power cell.³⁰ Like a breeding horse, an innovation is *fertile* if it can be used to create another innovation. Conversely, like a mule, an innovation is *infertile* if it cannot be used to create another innovation.

Economic concepts define “fertility” precisely: one innovation is *more* fertile than another if the costless transfer of resources from the latter to the former increases growth.³¹ By definition, a transfer is *costless* if the deadweight loss is nil—the loser loses the same amount as the winner wins. When transfers are costless, transferring resources from relatively infertile to relatively fertile innovations increases growth. To maximize growth, costless redistribution should proceed until all developed innovations are equally fertile.

The preceding proposition concerns costless transfers, whereas most transfers are costly. When transfers are costly, maximizing growth is more complicated. It requires balancing the deadweight cost of transfers and the benefits from funding more fertile inventions. The balance applies to the strength of patents. To maximize growth, a relatively fertile innovation's patent rights should be strengthened against others using it to innovate until the increase in growth from the more fertile innovation equals the decrease from more deadweight loss.³²

³⁰ Brian Arthur describes technology as “self-creating” (autopoietic). “[I]f new technologies were constructed from existing ones, then considered collectively, technology created itself.” W. BRIAN ARTHUR, *THE NATURE OF TECHNOLOGY* 2 (2009).

³¹ In a costless transfer from j to i , the burden of the transfer on j equals the benefit of the transfer to i . Mathematical notation makes this characteristic more precise. In notation, let $n \in \{0, 1, 2, \dots, N\}$ denote an innovation. The winner of an open competition to make innovation n enjoys profits denoted δ_n . The sustainable growth rate g is a function of the distribution of profits: $g = g(\pi_0, \pi_1, \pi_2, \dots, \pi_N)$. Further, assume that j is “economically more fertile” than i . Then a costless transfer of profits δ_n from i to j increases the growth rate:

$$g(\pi_0, \pi_1, \pi_2, \dots, \pi_N) < g(\pi_0, \pi_1, \pi_2, \pi_i - \delta_n, \dots, \pi_j + \delta_n, \dots, \pi_N).$$

³² Here is the fertility principle's mathematical expression. If innovation j is “economically more fertile” than innovation i , then a costless transfer of profits δ_n from i to j increases the growth rate: $g(\pi_0, \pi_1, \pi_2, \dots, \pi_N) < g(\pi_0, \pi_1, \pi_2, \pi_i - \delta_n, \dots, \pi_j + \delta_n, \dots, \pi_N)$. However, an actual transfer caused by law and policy creates a deadweight loss. The burden on i is δ_n , the transfer to j is $\gamma\delta_n$, and the deadweight loss is $(1 - \gamma)\delta_n$, where $\gamma \in [0, 1]$. At the optimum as given by the fertility principle, the benefit from the transfer to a more fertile innovation exactly offsets the cost from the deadweight loss, so the growth rate does not increase or decrease: $g(\pi_0, \pi_1, \pi_2, \dots, \pi_N) < g(\pi_0, \pi_1, \pi_2, \pi_i - \delta_n, \dots, \pi_j + \gamma\delta_n, \dots, \pi_N)$.

In sum, according to the fertility principle, *to promote innovation, patent protection should be strong for a very fertile innovation against an infertile innovation*. By “very” fertile, we mean fertile enough to overcome the cost of redistribution.

The separation and fertility principles are the foundations of patent law in growth economics. The separation principle is simpler to apply because it only requires distinguishing innovation from consumption and production, whereas the fertility principle requires balancing fertility and deadweight loss.³³ Applying these principles to patent laws can illuminate their effects, justification, and interpretation. In this Article, we apply the separation principle at length, and we mention the fertility principle briefly.³⁴

C. Welfare Overtaking

The separation and fertility principles predict how patent law can maximize the rate of innovation. This goal matters for implementing patent law’s constitutional purpose. Legality aside, what about patent law’s normative justification? What is the moral, philosophical, political, or policy justification of patent law? In economics, normative justification focuses on human welfare.

Robert Lucas, the economist who won the Nobel Prize in 1995, famously commented, “Once one starts to think about economic growth, it is hard to think about anything else.” Compared to sustained growth, other economic sources of human welfare are insignificant. Compounded over a century, 2% annual growth

³³ Vast literature on cumulative innovation attempted to shed light on the optimal strength of legal protection based on the degree of innovative contribution. We do not enter into this complicated debate because our focus here is on the separation principle. For insight into the literature on cumulative innovation, see, for example, Suzanne Scotchmer, *Standing on the Shoulders of Giants*, 5 J. ECON. PERSP. 29 (1991); Jerry R. Green & Suzanne Scotchmer, *On the Division of Profit in Sequential Innovation*, 26 RAND J. ECON. 20 (1995); James Bessen & Eric Maskin, *Sequential Innovation, Patents, and Imitation*, 40 RAND J. ECON. 611 (2009); John H. Barton, *Patents and Antitrust: A Rethinking in Light of Patent Breadth and Sequential Innovation*, 65 ANTITRUST L. J. 449, 453 (1997); Howard F. Chang, *Patent Scope, Antitrust Policy, and Cumulative Innovation*, 26 RAND J. ECON. 34 (1995); Ted O’Donoghue, *A Patentability Requirement for Sequential Innovation*, 29 RAND J. ECON. 654 (1998); Robert P. Merges & Richard R. Nelson, *On the Complex Economics of Patent Scope*, 90 COLUM. L. REV. 839, 876–79 (1990); Robert P. Merges, *Rent Control in the Patent District: Observation on the Grady-Alexander Thesis*, 78 VA. L. REV. 359 (1992) [hereinafter *Rent Control*]; Robert P. Merges, *Intellectual Property Rights and Bargaining Breakdown: The Case of Blocking Patents*, 62 TENN. L. REV. 75 (1994) [hereinafter *IP Bargaining Breakdown*]; Lemley, *supra* note 25; Kevin E. Collins, *Getting into the “Spirit” of Innovative Things: Looking to Complementary and Substitute Properties to Shape Patent Protection for Improvements*, 26 BERKELEY TECH. L.J. 1217 (2011).

³⁴ See *infra* Part V.

(roughly the growth rate of the U.S. economy over a century) increases wealth more than 7 times, and 10% annual growth (roughly the growth rate of the Chinese economy from 1980 to 2010) increases wealth by almost 14,000 times.³⁵ People underestimate the effects of compound growth because they forget that a 5% increase is much larger absolutely for a teenager than a toddler.

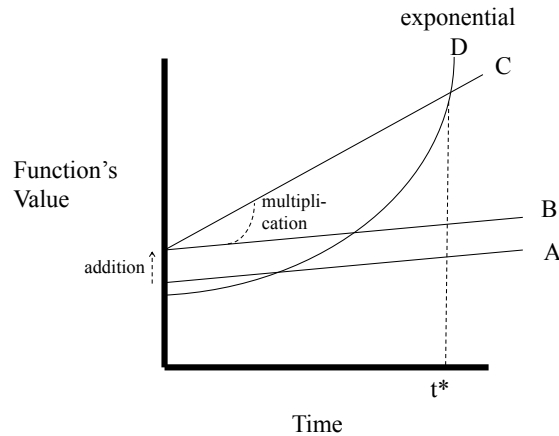
Behind these observations is a mathematical truth: an economy that increases at a constant proportional rate will overtake an economy that increases at a constant absolute rate. Figure 6 depicts this fact. The vertical axis in Figure 6 represents a function's value, and the horizontal axis represents time. Start with function A, whose value increases at a constant absolute rate with time, as indicated by A's constant slope. Consider changes to function A. An addition to the value of A at time 0 shifts A up and yields B. B starts at a higher level than A, and B grows at the same absolute rate as A. Next, multiply B by a constant value, which rotates the graph of B up to C. C grows at a higher absolute rate than B. Finally, contrast line C with curve D. At time 0, D starts below C, and then D grows at a faster proportional rate than C, overtaking C at time t^* . In general, a function with constant proportional growth like D will overtake a function with constant absolute growth like C.

³⁵ Here is a table of the size reached by an economy that starts at 1 and grows at various rates and years:

Growth rate	1 year	5 years	10 years	25 years	50 years	100 years
0%	1.00	1.00	1.00	1.00	1.00	1.00
1/2%	1.005	1.01	1.05	1.13	1.28	1.65
1%	1.01	1.05	1.10	1.28	1.64	2.70
2%	1.02	1.10	1.22	1.64	2.69	7.24
5%	1.05	1.28	1.63	3.39	11.47	131.50
10%	1.10	1.61	2.59	10.83	117.39	13,780.61

*size of economy = $(1 + r/100)^t$, where r = percentage growth rate, and t = years of growth.

Figure 6. Exponential Growth Overtakes Additive and Multiplicative Increases



Efficiency and innovation build on themselves differently. Increases in static efficiency cause one-time growth, as in the shift from A to B in Figure 6. Innovation sometimes causes growth that builds on itself at a constant absolute rate, as in the shift from B to C. Innovation sometimes causes growth at a constant proportional rate, as in D. The development in modern economies that has greatly—indeed, miraculously—increased human welfare is constant proportional growth, such as 2% in GDP. Relative to the past, the business cycle has dampened, and growth has proceeded quickly in many economies. Consequently, modern economies are usually compared in terms of proportional growth rates.

Now we apply Figure 6 to patent law. Think of function C as the path of social welfare when a certain class of innovations are unpatentable. Now, imagine a change in the law that makes these innovations patentable. When the new patents create market power for innovations, their prices increase for consumers. The increase in consumer prices reduces welfare. The reduction in welfare takes the form of an immediate loss, represented by the shift in Figure 6 from B down to A.³⁶ Thus, the difference between B and A represents the total loss in welfare from static inefficiency caused by patents.

Next, we interpret Figure 6 to represent the gain in welfare from the innovation. Assume that the law changes from no-patent to patent. The immediate effect is static inefficiency, as represented by a shift from B to A. The subsequent effect is proportional growth, as represented by the shift from A to D. Thus, the total effect is $B \rightarrow A \rightarrow D$. In sum, patent law causes the economy to shift from path B to path D. Prior to time t^* , B lies above D, which indicates that the loss in static efficiency exceeds the gain from innovation. At

³⁶ The immediate loss corresponds to the “welfare triangle” in the usual one-period graph of the loss in welfare from a monopoly price.

time t^* , the path D overtakes the path B. After that, the gain from innovation exceeds the loss in static efficiency by a steadily increasing amount. If overtaking is fast and D is growing quickly, then growth overshadows inefficiency, and static inefficiency is unimportant to total welfare. This is the overtaking principle.³⁷

To illustrate, consider software applications written for consumers. Patents on software applications make consumers lose immediately from higher prices and gain subsequently from higher performance. If overtaking is quick, the subsequent gain from better applications easily exceeds the immediate loss from higher prices. Overtaking justifies evaluating laws concerning their dynamic effects (e.g., the rate of growth), not their static effects (e.g., immediate increase in consumer prices). Consequently, the welfare effects of patents laws for computer software can be assessed based on rates of innovation.

Similarly, consider pharmaceutical drugs. Strong patents against generic drug manufacturers make consumers lose immediately from costlier drugs and gain later from better drugs. Legal commentators disagree sharply over the balance of these effects. In contrast, the overtaking principle offers a decisive answer. If overtaking is quick, the gain in welfare from better drugs will quickly overtake the loss from higher prices.³⁸

³⁷ A theorem along these lines is proved in Robert Cooter & Aaron Edlin, *Maximizing Growth vs. Static Efficiency or Redistribution* (Berkeley Law & Econ. Working Paper, 2010). Assume that utility in a generation is a function of its consumption, and social welfare equals a discounted sum of the utilities of the generations. Consider two sequences of consumption in an infinite number of generations. In the first sequence, initial consumption is higher, and growth is lower. In the second sequence, initial consumption is lower, and growth is higher. If the social welfare function is reasonably convex in the utility of the generations, and the utility function of each generation is reasonably convex in consumption, then the sum of utilities in the second sequence eventually overtakes the sum of utilities in the first sequence.

³⁸ The overtaking principle may be criticized on moral grounds—until overtaking is completed, present generations would bear the seemingly unfair price of static inefficiency, whereas future generations would enjoy the spoils of economic growth. Indeed, preventing concrete suffering from present generations may seem more ethically compelling than preventing speculative suffering (even if more substantial) from future generations. The Israeli government, for example, embraced this moral concern when it decided to release one captive Israeli soldier in exchange of over 1,000 Palestinian prisoners, many of them convicted terrorists, despite repeated warnings by the Israeli intelligence that this exchange deal would lead the death of many future Israeli civilians. See Ronen Bergman, *Gilad Shalit and the Rising Price of an Israeli Life*, N.Y. TIMES (November 9, 2011), <https://www.nytimes.com/2011/11/13/magazine/gilad-shalit-and-the-cost-of-an-israeli-life.html>. In the patent sphere, such arguments are often heard in the area of pharmaceuticals where patients of low socioeconomic status

When comparing the welfare effects of one patent law to another, the overtaking principle reduces the comparison to growth rates. According to the overtaking principle, a law is optimal when it maximizes the rate of innovation. In contrast, conventional analysis of patents balances growth and efficiency. The balancing test compares the immediate costs of inefficient resource allocation from a patent and the present value of future benefits from innovation. Present value calculations involve controversial assumptions about discounting the future. Also, combining the benefits and costs of different people involves controversial assumptions about the marginal utility of income. Consequently, present value is much more complicated than overtaking when evaluating patent law.

Applying the overtaking principle requires measuring growth. Economists have a long tradition of measuring growth in different ways. Almost everyone agrees that income, consumption, wealth, and other economic measures should be regarded as surrogates for something else that is intrinsically valuable. However, people disagree about what is intrinsically valuable. Thus, economists regard “welfare” as the intrinsic value behind income and wealth, but they disagree about how to measure it.

Fortunately, the overtaking principle is robust with respect to different measures of growth. Its usefulness in patent law does not rest on a particular measure of “progress” in the useful arts. For example, if stronger patents induce faster technical progress in pharmaceutical drugs, the gains will usually overtake the losses in static efficiency, regardless of whether gains and losses are measured by welfare, income, or wealth. Therefore, the overtaking principle can be applied without first solving some difficult ethical and philosophical disputes.

Reinterpreting Figure 6 illustrates the robustness of overtaking. Our preceding discussion interprets functions A, B, C, and D as representing welfare. Instead of welfare, the functions in Figure 6 might represent, say, per capita income, consumption, or wealth. Thus, Figure 6 can be interpreted as showing an exponential increase in, say, income per capita, overtakes an additive or

struggle to meet the high price-tag of lifesavings drugs. *See generally* Hacothen, *supra* note 14 (discussing the social implication of overly prolonged drug monopolies). These considerations are unquestionably important, but we do not think they are part of the constitutional patent mandate. Moral and humanitarian considerations should be accounted for either through welfare policy or, in specific areas of innovation by non-patent rewarding mechanisms. Various commenters, for example, have called for using subsidies and prizes to incentivize lifesaving pharmaceuticals instead of patents. *See generally* WILLIAM W. FISHER III & TALHA SYED, *INFECTION: THE HEALTH CRISIS IN THE DEVELOPING WORLD AND WHAT WE SHOULD DO ABOUT IT*, ch. 2 (forthcoming, Stanford Univ. Press).

multiplicative increase. A change in interpretation of the functions does not change the fact that growth overtakes static efficiency.

To be more precise, different interpretations of the functions do not affect the existence of an overtaking point t^* , but they do affect how long overtaking takes or the value of t^* . For example, income often correlates with welfare. Faster growth in one goes with faster growth in the other. Consequently, maximizing the growth of income usually maximizes growth in welfare. This fact justifies comparing patent laws with respect to growth in income, which is relatively easy to measure, instead of growth in welfare, which is relatively difficult to measure.³⁹

The amount of welfare created by a given income depends on who gets it and how it is spent. Most people agree that a more equal distribution of income increases welfare, although they disagree about moral and political philosophies. In economic terms, they believe an additional dollar for the poor increases welfare more than an additional dollar for the rich, although they disagree about how much more.⁴⁰ In an exponentially growing economy, however, growth overshadows income redistribution with respect to welfare. If overtaking occurs quickly, then equality is important to welfare

³⁹ Economists assume a causal connection between income and welfare as formalized through the “social welfare function.” In this formalization, the welfare of society is a function of the utility of the individuals in it, the utility of individuals is a function of the goods that they consume, and the goods that they consume is a function of their income. Thus, the value of the social welfare function increases as individual incomes increase. More income for individuals causes them to consume more, which increases individual utility, which increase social welfare. The concept of social welfare as a function of individual utility was introduced by Abram Bergson, *A Reformulation of Certain Aspects of Welfare Economics*, 52 Q. J. ECON. 310 (1938). For a recent discussion of social welfare that is deep but challenging to read, see MATTHEW D. ADLER, *WELL-BEING AND FAIR DISTRIBUTION: BEYOND COST-BENEFIT ANALYSIS* (Oxford Univ. Press 2011).

⁴⁰ According to an old tradition in economics, the “marginal utility of money” is higher for the poor because they use money to satisfy more urgent needs than the rich, as when the poor buy bread and the rich buy cake. Consequently, redistribution from the rich to the poor increases welfare directly. In Adler’s theory, social welfare increases with individual wealth, and social welfare increases with a more equal distribution of wealth or utility across individuals. ADLER, *supra* note 39. However, economists and philosophers dispute the rate at which marginal utility declines with income. Egalitarians believe that poor people get much more welfare from additional income than rich people, whereas anti-egalitarians believe that poor people get little or no more welfare from additional income than rich people. Thus, anti-egalitarians may believe that a growth of 4% in income causes similar growth in welfare, whereas egalitarians believe that a growth of 4% in income causes slower growth in welfare. Consequently, anti-egalitarians should believe that an income growth of 4% will cause social welfare to overtake the static effects quickly, whereas egalitarians should believe that overtaking will take longer.

for its effects on growth, not as an end in itself. For the welfare of the poor, the important question about equality is whether it increases or decreases the growth in their incomes.

To illustrate historically, the Cultural Revolution in China, which commenced in the 1960s and expired in 1975, pursued strict equality and destroyed growth. Reversing after 1980, China allowed innovators to keep much of the wealth they created. Inequality increased, economic growth exploded, and poverty plummeted. The lowest wage earners in China benefited far more from faster growth after 1980 than from more equality in the 1970s. The opposite can also occur. Too much inequality can slow growth, especially by undermining the health and education of workers. Thus, improved education and health of poor people would assuredly increase their productivity in, say, the U.S.A. or the Philippines.

Overtaking applies equally to efficiency and distribution. Static efficiency and income redistribution should not be regarded as the ends of patent law in a growing economy, but, rather, as means to increase growth.⁴¹ Overtaking, therefore, provides the normative justification for the separation and fertility principles.

As overtaking time decreases (i.e., as t^* falls), disagreements seem unimportant over the marginal utility of money or the discount rate for futurity. Once you start thinking about welfare overtaking, you care more about the effect of static efficiency and income redistribution on growth, and less about these elements for their own sake. For example, the division of the surplus from cooperation poses one of the most difficult disputes in ethics and political philosophy. Framed in a static economy, this seems like an insolvable problem of distributive justice. Framed in an exponential economy, this seems like a manageable problem of economic theory and statistical evidence. When welfare grows, scarcity fades, and so should disputes about distributive justice.

For example, life expectancy at birth is 82 years in Switzerland and 69 years in Bangladesh,⁴² and enrollment in secondary school is 96 percent among Swiss children of the appropriate age and 46 percent in Bangladesh.⁴³ Facts like these

⁴¹ Specifically, static efficiency gets the prices of goods aligned with their value to people. When prices align with value, innovator aim growth at the right objectives—increases in goods as valued by consumers. In brief, efficiency contributes to welfare by aiming growth.

⁴² *Life Expectancy at Birth, Total (Years)*, WORLD BANK, <https://data.worldbank.org/indicator/sp.dyn.le00.in?end=2008&start=1960> (last visited Feb. 9, 2020) (data for the year of 2008). Note that life expectancy is in the 40s in many African countries.

⁴³ *School Enrollment, Secondary (% gross)*, WORLD BANK, <https://data.worldbank.org/indicator/SE.SEC.ENRR?end=2008&start=1970&view=chart> (last visited Feb. 9, 2020) (data for 2007–08).

make almost everyone agree that human welfare is higher in Switzerland than Bangladesh, primarily because Switzerland is much wealthier than Bangladesh. They can agree with this conclusion, even though they disagree about welfare's meaning and measure. Consequently, they should agree about favoring laws that increase growth, even though they disagree about, say, the marginal utility of income and the discount rate for futurity. Equalizing incomes in Bangladesh seems less urgent than raising welfare to a level similar to Switzerland.

IV. OPTIMAL STRENGTH OF PATENTS

The patent for inventing a machine, creating a molecule, extracting a vaccine, or writing a computer program can be weak or strong. Much analysis of patent law concerns strength, of which three attributes are duration, breadth, and remedy. To maximize innovation, these attributes should separate consuming and producing from innovating. For inventions mostly used for producing and consuming, strong patents will transfer more wealth from static to dynamic uses, so patents should be long, broad, and powerfully remedied. Conversely, for inventions mostly used for innovating, strong patents burden innovators with a deadweight loss from redistribution, so patents should be short, narrow, and weakly remedied.

In this section we apply the separation principle to the rules that govern patents' duration, breadth, and remedy. We highlight instances where these rules confirm with the separation principle and instances where they do not. Later, in Section V.A we take a closer look at how specific doctrines should be tailored to better adhere to the separation mandate.

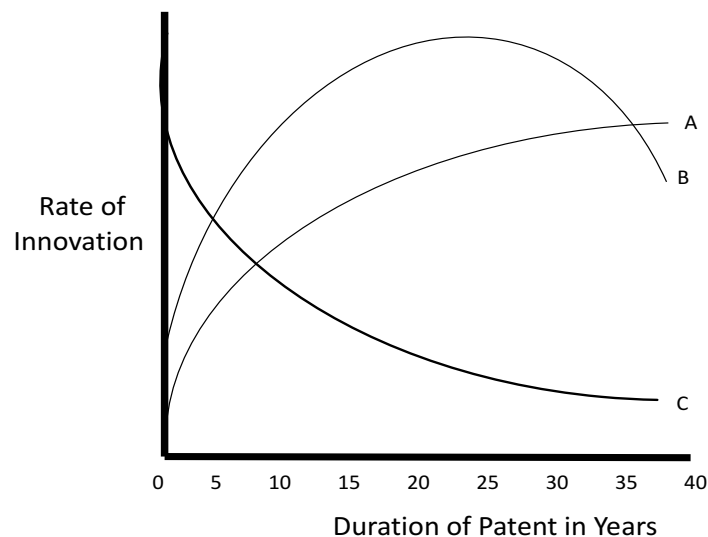
A. Duration

We begin with the optimal duration of patents. A patent's legal duration is usually defined by a statute and measured in years. The legal duration of patents affects their economic value, and their value affects the rate of innovation. Thus, the horizontal axis in Figure 7 indicates the length of patents in years, and the vertical axis indicates the rate of innovation. The curves in Figure 7 depict some possible relationships between the rate of innovation and patent length.

For curve A, the rate of innovation is an increasing function of patent length, so the innovation-maximizing patent length is infinite. Curve A might depict an industry where the optimal patent is very strong. According to the separation principle, patents should

be long against consuming and producing. Thus, pharmaceutical drugs sold for consumption or production might have a profile like Curve A.⁴⁴ Indeed, pharmaceutical innovation is usually discrete, not cumulative, and thus has limited impact on follow-on innovation.⁴⁵

Figure 7. Rate of Innovation and Patent Length



⁴⁴ See, e.g., JAMES BESSEN & MICHAEL J. MEURER, PATENT FAILURE: HOW JUDGES, BUREAUCRATS, AND LAYERS PUT INNOVATORS AT RISK 88–89 (2008) (noting the great importance of patents in the pharmaceutical industry in comparison to most other industries); ADAM B. JAFFE & JOSH LERNER, INNOVATION AND ITS DISCONTENTS: HOW OUR BROKEN PATENT SYSTEM IS ENDANGERING INNOVATION AND PROGRESS, AND WHAT TO DO ABOUT IT 39–41 (2004); Edwin Mansfield, *Patents and Innovation: An Empirical Study*, 32 MGMT. SCI. 173, 174 (1986); Richard C. Levin et al., *Appropriating the Returns from Industrial Research and Development*, 1987 BROOKINGS PAPERS ECON. ACTIVITY 783, 796, 824; and Jay P. Kesan, *Transferring Innovation*, 77 FORDHAM L. REV. 2169, 2195 (2009).

⁴⁵ See Hacothen, *supra* note 14 (explaining that many follow-on pharmaceutical improvements have marginal social value); Merges & Nelson, *supra* note 33, at 880 (1990) (defining discrete invention as an invention is “well-defined, created through the inventor’s insight and hard work . . . [and] does not point the way to wide ranging subsequent technical advances.” And arguing that “many new pharmaceuticals may [] fit this model.”); Olga Gurgula, *Strategic Accumulation of Patents in the Pharmaceutical Industry and Patent Thickets in Complex Technologies—Two Different Concepts Sharing Similar Features*, 48 INT’L REV. INTELL. PROP. & COMPETITION L. 385, 393 (2017) (“the nature of the technology—most often, it is discrete”); and BURK & LEMLEY, *supra* note 9, at 81 (“[P]harmaceutical patents do not merely cover small components that must be integrated into a marketable product, and this in turn means that a company that wishes to sell a pharmaceutical product generally won’t need licenses for many different patents.”).

Now consider the symmetrical opposite of Curve A. For Curve C, the rate of innovation is a decreasing function of patent length, so the innovation-maximizing patent length is zero. A profile much like Curve C justifies removing legal constraints for a class of inventions. According to the separation principle, patents should be short against innovating. In contrast to pharmaceutical innovation, software is a cumulative process done incrementally by many innovative contributors.⁴⁶ As such, software patents allegedly deserve a shorter term of patent protection,⁴⁷ or even no patent protection at all as advocated by the Free Software Foundation.⁴⁸

Finally, Curve B depicts the intermediate case in between Curve A and Curve C. Starting with short patents to the left on Curve B, moving to the right and lengthening the patent increases the rate of innovation at first. A point is reached, however, where further lengthening decreases innovation. For Curve B, the innovation-maximizing patent length is intermediate, not infinite or zero. In the intermediate case, moderately strong intellectual property rights maximize the rate of innovation.⁴⁹

The international standard for duration, which the U.S. has followed, is 20 years from the date of the application for a patent.⁵⁰ As depicted in Curve B, the optimal patent duration is 20 years. Thus, the international standard would maximize innovation if the typical patent profile actually resembles Curve B, and the highest point occurred at 20 years.

The profiles in Figure 7 could be interpreted as describing three different types of innovations. Averaging them would yield a single standard. Instead of a single standard, a more faithful

⁴⁶ Peter S. Menell, *Tailoring Legal Protection for Computer Software*, 39 STAN. L. REV. 1329, 1338 (1986) (noting that with respect to software development, “secondary inventions—including essential design improvements, refinements, and adaptations to a variety of uses—are often as crucial to the generation of social benefits as the initial discovery”).

⁴⁷ *Id.* at 1364 (offering that legal protection for software should be short in duration).

⁴⁸ See *What is Free Software?*, GNU, <https://www.gnu.org/philosophy/free-sw.html.en> (last updated July 30, 2019) (explaining that users must have “the freedom to run, copy, distribute, study, change and improve the software.”).

⁴⁹ Note that Curves A, B, and C all have the same rule to find the maximum: starting from 0, increase the duration of intellectual property rights so long as venture profits increase. Stop increasing the duration intellectual property rights when venture profits stop increasing. For analyzing many activities, economists assume a concave function with a unique maximum. Another possibility, which we will not discuss, is a function with several local maxima.

⁵⁰ 35 U.S.C. § 154 (2018). Before June 8, 1995, U.S. patent duration was 17 years from the date the patent issued. This was changed to comply with the international standard. See Agreement on Trade-Related Aspects of Intellectual Property Rights, Apr. 15, 1994, art. 33, 33 I.L.M. 81 (1994).

application of the separation principle would require different patent lengths applied in different industries, such as pharmaceuticals, machine tools, and software.⁵¹ How long should patent protection be for each innovative industry? Existing answers are intuitive, not scientific. No compelling empirical evidence indicates whether the pace of innovation would increase from changing the duration of patents. Quantitative research on the optimal duration of intellectual property rights has barely begun.⁵²

In any case, research on patent duration should apply the separation principle, not overlook it. Thus, when Landes and Posner advocate patents of infinite length, they do not distinguish between patents against innovating and patents against consuming and producing.⁵³ According to the separation principle, the former should be longer than the latter. In a similar vein, while many foreign jurisdictions issue “petty” patents of short duration alongside “full” patents of longer duration, these authorities do not systematically separate innovating from consuming or producing.⁵⁴

B. Breadth

Besides duration, patents have breadth. To illustrate breadth, the large circle in Figure 8 depicts the space of possible innovations, with the line R bisecting it into “ownable” and “unownable.”⁵⁵ Ownable subject matter includes inventions such as computer chips,

⁵¹ See generally BURK & LEMLEY, *supra* note 9 (exploring industry-based patent levers). Unlike breadth or remedy that are subject to some degree to the court’s equitable discretion, changing the duration of patents would require a legislative action. Nevertheless, by tailoring patent’s breadth and remedy, courts can shorten the “effective duration” of patents in term of their market impact. For example, tailoring patents’ breadth to facilitate an earlier development of a non-infringing substitute would effectively shorten the patent monopoly. See Ted O’Donoghue, Suzanne Scotchmer & Jacques-Francois Thisse, *Patent Breadth, Patent Life, and the Pace of Technological Progress*, 7 J. ECON. MGMT. STRATEGY 1, 1 (1998) (“[A] patent can terminate either because it expires or because a non-infringing innovation displaces its product in the market.”).

⁵² See Eric Budish, Benjamin N. Roin & Heidi L. Williams, *Patents and Research Investments: Assessing the Empirical Evidence*, 106 AM. ECON. REV. 1, 1 (2016) (describing the body of empirical evidence on this issue as “surprisingly small”).

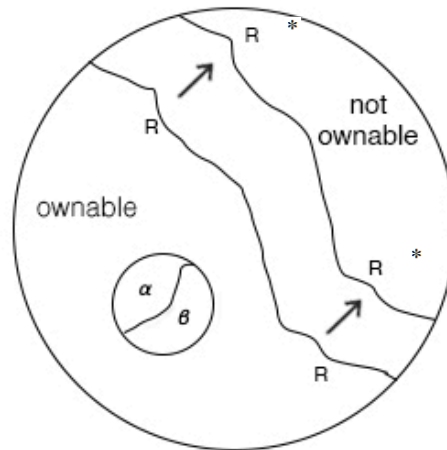
⁵³ RICHARD A. POSNER & WILLIAM M. LANDES, *THE POLITICAL ECONOMY OF INTELLECTUAL PROPERTY LAW* 294 (Belknap Press 1st ed. 2003) (advocating for infinite patents as a way of simplifying property rights and lubricating bargaining in all industries).

⁵⁴ See generally *Utility Models*, WORLD INTELL. PROP. ORG., https://www.wipo.int/patents/en/topics/utility_models.html (last visited Aug. 16, 2019); *Utility Model Protection*, GER. PAT. & TRADEMARK OFF. (DPMA), https://www.dpma.de/english/utility_models/utility_model_protection/index.htm 1 (last visited Aug. 6, 2019).

⁵⁵ This is a picture, not a graph, because the breadth of innovations has no natural measure. See *infra* note 74 (discussing the difference between an ordering and a measure).

pharmaceutical drugs, and machine screws. Unownable subject matter includes discoveries that are not applied inventions and abstract ideas.⁵⁶

Figure 8. The Breadth of Patent Protection



In Figure 8, the boundary between ownable and unownable innovations represents the reach of the patent system as a whole. Expanding the reach of the patent system shifts the boundary from R to R* in Figure 8. For example, U.S. patent law increased its reach when inventors were allowed to patent business processes for the first time, such as Amazon’s “one-click” ordering from its online catalog.⁵⁷ Expanding patent’s reach fences in more of the common land of innovation.⁵⁸

⁵⁶ See *infra* Section V.A.i.

⁵⁷ U.S. Patent No. 5,960,411A (filed Sept. 12, 1997). *State Street Bank & Trust v. Signature Financial Group*, 149 F.3d 1368 (Fed. Cir. 1998), is considered to be the decision that opened the floodgate of business method and software patents. See *Hughes Aircraft Co. v. United States*, 148 F.3d 1384, 1385 (Fed. Cir. 1998) (Clevenger, J., dissenting) (observing that, after *State Street*, “virtually anything is patentable”); see also William T. Ellis & Aaron C. Chatterjee, *State Street Sets Seismic Precedent*, NAT’L L.J., Sept. 21, 1998, at B13 (advising financial institutions how to react to *State Street*); Barry D. Rein, *A New World for Money Managers*, N.Y.L.J., Sept. 21, 1998, at S1 (explaining that industry, before *State Street*, generally disregarded patents and assumed that financial software was not within the patent system); Teresa Riordan, *An Appeals Court Says a Mathematical Formula Can Be Patented, If It Is a Moneymaker*, N.Y. TIMES, Aug. 3, 1998, at D4 (reporting that *State Street* went further than expected).

⁵⁸ In our terminology, we will use “invention,” and “innovation” interchangeably to describe subject matter that lies within the reach of the patent system (satisfies the requirements of eligibility and utility as we shall describe). We will avoid the confusing distinction between an “invention,” and an “innovation” that derived from the Schumpeterian tradition. JOSEPH A. SCHUMPETER, *CAPITALISM, SOCIALISM, AND DEMOCRACY* 106 (6th ed. 1987); see also RICHARD R. NELSON

The separation principle mandates that patents should be weak or nonexistent against innovation. Accordingly, “abstract ideas, laws of nature, and physical phenomena” that are fundamental for spurring future innovation might as well be excluded from patent protection altogether.⁵⁹ Indeed, in a line of cases dating back to the early 1850s, courts have systematically singled out these categories as judicially made exceptions to patent eligibility and limited the reach of the patent system.⁶⁰ Beginning with *Gottschalk v. Benson*,⁶¹ the Supreme Court even grounded its intuition for making these judicial exceptions in the separation principle by saying that legal protection over “the basic tools of scientific and technological work,” would “preempt” future innovation by others.⁶²

Alongside the doctrine of patent eligibility, courts have also used the utility doctrine to narrow the reach of the patent system in accordance with the separation principle.⁶³ For example, the utility requirement was interpreted by courts to exclude from patent protection short DNA molecules that have no use in production and consumption but are valuable for research.⁶⁴

& SIDNEY G. WINTER, AN EVOLUTIONARY THEORY OF ECONOMIC CHANGE 263 (1982) (distinguishing the invention of a product from innovation, a broader process of research, development, testing, and commercialization of that product, and attributing that distinction to Schumpeter); WILLIAM KINGSTON, DIRECT PROTECTION OF INNOVATION 13 (William Kingston ed., 1987).

⁵⁹ See *Diamond v. Chakrabarty*, 447 U.S. 303 (1980).

⁶⁰ See *Le Roy v. Tatham*, 55 U.S. (14 How.) 156, 174–75 (1852) (“A principle, in the abstract, is a fundamental truth; an original cause; a motive; these cannot be patented, as no one can claim in either of them an exclusive right.”).

⁶¹ 409 U.S. 63 (1972)

⁶² *Id.* at 68, 71–72 (“The mathematical formula involved here has no substantial practical application except in connection with a digital computer, which means that if the judgment below is affirmed, the patent would *wholly pre-empt* the mathematical formula and in practical effect would be a patent on the algorithm itself.” (emphasis added)). The Supreme Court used the term “preemption” more recently in *Bilski v. Kappos*, 561 U.S. 593 (2010). Since then, the use of this term caused endless confusion. See Pamela Samuelson, *Benson Revisited: The Case Against Patent Protection for Algorithms and Other Computer Program-Related Inventions*, 39 EMORY L.J. 1025 (1990); Rochelle C. Dreyfuss & James P. Evans, *From Bilski Back to Benson: Preemption, Inventing Around, and the Case of Genetic Diagnostics*, 63 STAN. L. REV. 1349, 1352 (2011).

⁶³ Both the doctrine of patent eligibility and of utility are based on 35 U.S.C. § 101 (1994). See *Kewanee Oil Co. v. Bicron Corp.*, 416 U.S. 470, 483 (1974) (holding that inventions cannot receive patent protection unless they fall within express category of 35 U.S.C. § 101).

⁶⁴ See Utility Examination Guidelines, 66 Fed. Reg. 1092, 1098 (Jan. 5, 2001) (for patents drawn to such molecules requiring a showing of “specific,” “substantial,” and “credible” application for the technology); *In re Fisher*, 421 F.3d 1365 (Fed. Cir. 2005). According to the separation principle, the legal protection for research tools should indeed be narrow (as these inventions are used primarily for innovation). Nevertheless, this does not necessarily mean categorical exclusion. See Section V.A.

While the logic of these judicial exclusions is sound, the doctrines that govern the reach of the patent system are binary and aggressive and should be used in caution.⁶⁵ In a series of recent decisions—most notably in *Mayo*⁶⁶ and *Alice*⁶⁷—the Supreme Court has overly expanded the judicial exceptions to patent eligibility beyond the spirit of the separation principle.⁶⁸ In Part IV.B we explore the misleading narratives that led to these decisions. Following these precedents, it became exceedingly difficult to obtain patents in various cutting-edge, innovative industries, such as biotechnology and software.⁶⁹ Pending legislation seek to remove these barriers and thus broaden the reach of the patent system again.⁷⁰

Instead of narrowing the patent system’s reach, patent policy could adhere to the separation principle more subtly, by tailoring the scope of patents.⁷¹ A patent’s scope refers to all inventions that infringe on a patent if practiced without a license.⁷² To represent

⁶⁵ Indeed, in all industries other than biotechnology and chemistry the utility requirement was interpreted extremely narrowly. *See Juicy Whip, Inc. v. Orange Bang, Inc.*, 185 F.3d 1364, 1366 (Fed. Cir. 1999) (“The threshold of utility is not high.”); *Process Control Corp. v. HydReclaim Corp.*, 190 F.3d 1350, 1358 (Fed. Cir. 1999) (“[T]he subject matter of the claim must be operable.”). Also, for eligibility, early court decisions narrowly applied the judicial exclusions to principles in the abstract as oppose to technological application of such principles. *See Tatham*, 55 U.S. at 175; *see also* Brief of Professors Jeffrey A. Lefstin & Peter S. Menell as Amici Curiae in Support of Petitioner for a Writ of Certiorari at 4–14, *Sequenom, Inc. v. Ariosa Diagnostics, Inc.*, 136 S. Ct. 2511 (2016) (No. 15-1182).

⁶⁶ *Mayo Collaborative Servs. v. Prometheus Labs., Inc.*, 566 U.S. 66, 71 (2012).

⁶⁷ *Alice Corp. v. CLS Bank Int’l*, 573 U.S. 208 (2014).

⁶⁸ Some commenters have argued that the *Alice* decision allegedly “invalidate[d] the majority of all software patents in force today.” *See CLS Bank Ruling a Big Deal for Valley Software Patents*, STAN. L. SCH. (June 19, 2014), <https://law.stanford.edu/press/cls-bank-ruling-a-big-deal-for-valley-software-patents> (statement by professor Mark Lemley to The Recorder). On the challenges that the *Mayo*–*Alice* framework has caused various industries, *see* Jeffrey A. Lefstin, Peter S. Menell & David O. Taylor, *Final Report of the Berkeley Center for Law & Technology Section 101 Workshop: Addressing Patent Eligibility Challenges*, 33 BERKELEY TECH. L.J. 551 (2018).

⁶⁹ *See* Lefstin, Menell & Taylor, *supra* note 68, at 581.

⁷⁰ *See* Press Release, *Sens. Tillis and Coons and Reps. Collins, Johnson, and Stivers Release Draft Bill Text to Reform Section 101 of the Patent Act*, THOM TILLIS, U.S. SENATOR FOR N.C. (May 22, 2019), *available at* <https://www.tillis.senate.gov/2019/5/sens-tillis-and-coons-and-reps-collins-johnson-and-stivers-release-draft-bill-text-to-reform-section-101-of-the-patent-act>.

⁷¹ *See* Dreyfuss & Evans, *supra* note 62 (distinguishing the “whether” inquiry, what we call “reach,” from the “how” inquiry, what we call “scope.”).

⁷² Janet Freilich, *The Uninformed Topography of Patent Scope*, 19 STAN. TECH. L. REV. 150, 195 (2015) (noting that patent scope “manifests in practice as the universe of inventions that infringe on the patent.”); *See also* *Festo Corp. v.*

scope graphically, we subdivide the small circle of inventions in Figure 8 into innovation α and innovation β . Assume that firm A invents α and firm B subsequently invents β . A claims that α and β are a single invention, and A files for a patent covering both of them. Later, B claims that α and β are different inventions, and B files for a patent for β . The patent office must decide whether A's patent precludes B from patenting β . For a given space of inventions, the state can generally give few broad patents or many narrow patents, rather like a city planner can divide undeveloped land into a few large lots or many small lots.⁷³

The concept of patent scope is difficult to grasp because scope has no simple measure. Thus, a patent covering rain gear is broader than a patent covering umbrellas; the former subsumes the latter. Given subsumption, we can say which patent is larger but not how much larger it is.⁷⁴ Subsumption orders scope without measuring it. Also, while subsumption orders inventions of the same type, it does not order inventions of different types, like an umbrella and an eggbeater. Laws and doctrines governing patent scope are

Shoketsu Kinzoku Kogyo Kabushiki Co., 234 F.3d 558, 589 (Fed. Cir. 2000) (Rader, J., concurring in part) ("The limitations of a patent's claims provide an initial measure of the effective scope of the patent. . . ."); Christopher A. Cotropia, *After-Arising Technologies and Tailoring Patent Scope*, 61 N.Y.U. ANN. SURV. AM. L. 151, 172–73 (2005) ("[T]he broader the patent scope, the more protection the patent holder receives and the more competing products she can exclude."); Merges & Nelson, *supra* note 33, at 839 ("The economic significance of a patent depends on its scope: the broader the scope, the larger the number of competing products and processes that will infringe on the patent"); *Smith v. Mid-Continent Inv. Co.*, 106 F.2d 622, 624 (8th Cir. 1939) ("[T]he scope of a patent is the boundaries (or limits) of the invention protected by the patent.").⁷³ Once lots are surveyed in a new town, owners can sell a lot but not half of a lot. The owner who cannot sell half of a lot can lease it for use by someone else. Much the same is true for selling patents. An owner cannot sell half of a patent, but an owner can contract for someone else to use half of it. Some theorists think that divisibility is the essential difference between property rights and contract rights. See Thomas W. Merrill & Henry E. Smith, *Optimal Standardization in the Law of Property: The Numerus Clausus Principle*, 110 YALE L. J. 1 (2000).

⁷⁴ Patent scope resembles the alphabet. Letters A through N encompass more letters than A through M, but there is no natural measure of the distance between M and N. Similarly, no natural measure exists to answer the question, "Is the difference in scope between a patent on rain gear and a patent on umbrellas larger or smaller than the difference in scope between a patent on umbrellas and a patent on automatically opening umbrellas?" In this respect, patent scope contrasts with height. One person is taller than another and the difference can be measured in centimeters. Fortunately, an ordering without a measure can support the mathematics of maximization. This fact makes modern utility theory possible in economics. Economic models often maximize utility functions that order states of the world by preference ("ordinal utility"), but do not measure differences in utility levels. Similarly, finding the scope of patent that maximizes growth only requires an ordering by scope, not a measure.

necessarily complicated because breadth has many immeasurable dimensions.⁷⁵

Various legal doctrines set patent scope during patent examination and again during litigation. When filing for a patent, the applicant makes claims about the innovation's scope. Most claims can be ordered from broad to the narrow. The United State Patent and Trademark Office (PTO) rejects or authorizes the claims in light of the prior art⁷⁶ and the richness of the applicant discloser.⁷⁷ The issuance of a patent implies a legally permissible scope for a patent.⁷⁸

Then, the patent scope is scrutinized once again in litigation. The accused infringer would often allege that the issued patent is overbroad and should either be invalidated or narrowed to exclude the allegedly infringing use.⁷⁹ A court would scrutinize the PTO's decision when issuing the patent, based on the information that was available to the office when the patent was prosecuted.⁸⁰ The court would also consider information that became available post-prosecution by the time of the litigation, such as the nature of the allegedly infringing use, and would tailor patent scope accordingly.⁸¹ Many doctrines operate during the litigation phase, allowing courts to tailor patent scope.⁸² These doctrines including

⁷⁵ See generally Merges & Nelson, *supra* note 33; Freilich, *supra* note 72.

⁷⁶ 35 U.S.C. §§ 102–103 (2018) (disallowing claims for known or obvious subject matter).

⁷⁷ 35 U.S.C. § 112(a) (2018) (limiting applications covered by a general claim to those that a person of ordinary skill in the art could make without undue experimentation (“enablement”), and to subject matter that was in the inventor’s possession in the time of the application (“written description.”)). See U.S. PATENT & TRADEMARK OFFICE, MANUAL OF PATENT EXAMINING PROCEDURE ch. 2100, §§ 2163–64 (2018), <https://www.uspto.gov/web/offices/pac/mpep/mpep-2100.html>.

⁷⁸ 35 U.S.C. § 282(a) (“[A] patent shall be presumed valid.”).

⁷⁹ Other commenters also defined litigation-stage doctrines as levers that serve to tailor a patent’s scope toward the optimum. See, e.g., Freilich, *supra* note 72; Kevin Collins, *An Initial Comment on Ariad: Written Description and the Baseline of Patent Protection for After-Arising Technology*, 2010 PATENTLY-O PAT. L.J. 24 (defining “retrospective scope” for prior art doctrines, and “prospective scope” for doctrines that tailor the scope of the patents in view of technological advances that were made after the patent was filed).

⁸⁰ Indeed, when the court investigates the prior art and the disclosure doctrines, it evaluates them based on the information that was available before the PTO. Cf. Freilich, *supra* note 72, at 167 (calling these doctrines “ex ante tools for measuring patent scope”).

⁸¹ See Merges & Nelson, *supra* note 33, at 909 (“Once a court completes its assessment of the significance of the patented device, it should consider in addition the importance of the advance represented in the accused device.”).

⁸² Cf. Freilich, *supra* note 72, at 172 (calling these doctrines “ex post tools for measuring patent scope.”).

equivalence and reverse equivalence,⁸³ the rules of claim construction,⁸⁴ and numerous exceptions to infringement liability.⁸⁵

Within this territory, the separation reign is the strongest. Courts should weaken a patent's scope in light of innovative uses and sustain broader scope in light of uses in production and consumption. In a seminal paper, Robert Merges and Richard Nelson urged courts to consider as part of the infringement analysis how the invention was used rather than just the invention itself.⁸⁶ In close to the three decades that have passed since Merges and Nelson's recommendation, many legal developments have pushed to the opposite direction. In Part V, we review these troubling trends and explore how the separation principle could be applied more accurately.

⁸³ These doctrines broaden the scope of the claims to include minor alterations and narrows the scope to exclude radical alterations. The doctrine of equivalence was first introduced in *Winans v. Denmead*, 56 U.S. (15 How.) 330 (1853), but since then, it has been significantly refined. See *Graver Tank & Mfg. Co. v. Linde Air Prods. Co.*, 339 U.S. 605, 607–08 (1950). The reverse doctrine of equivalence dates back to the Supreme Court decision in *Westinghouse v. Boyden Power Brake Co.*, 170 U.S. 537 (1898), but since then, the doctrine was heavily beaten. See, e.g., *Tate Access Floors v. Interface Architectural Research*, 279 F.3d 1357, 1368 (Fed. Cir. 2002) (calling the doctrine an “anachronistic exception, long mentioned but rarely applied”). See *infra* note 153 and accompanying text.

⁸⁴ See Dan L. Burk & Mark A. Lemley, *Quantum Patent Mechanics*, 9 LEWIS & CLARK L. REV. 29, 29 (2005) (“Define an element narrowly—limit it to a single word, say—and you will tend to narrow the resulting patent. By contrast, defining an element broadly tends to broaden the patent.”); Peter S. Menell, Matthew D. Powers & Steven C. Carlson, *Patent Claim Construction: A Modern Synthesis and Structured Framework*, 25 BERKELEY TECH. L.J. 711, 742 (2010) (explaining that some doctrines of claim construction “tend to narrow claim scope, while others broaden it”).

⁸⁵ An allegedly infringing use can be liberated from infringement for many reasons. One reason can be the nature of the *contested use* (i.e., the experimental use defense can liberate innovative uses from infringement. See *infra* Section V.A.ii). Another reason may be the nature of *the user* (i.e., “shop rights,” which is a common law doctrine allowing an employer to use the employee's invention without payment if that invention was made using the employer's time, materials, facilities, or equipment. See, e.g., *Lariscey v. United States*, 949 F. 2d 1137 (Fed. Cir. 1991)). A third reason for liberating infringement may be the *improper behavior by the patent owner* (i.e. exhaustion, misuse, inequitable conduct, laches, estoppel etc.). The reverse doctrine of equivalence is also a defense of patent infringement. Nevertheless we have classified it within the general doctrines that govern patent scope because unlike the other defenses it can be viewed as excluding subject matter from the scope of the original patent, rather than merely excusing the infringement.

⁸⁶ See Merges & Nelson, *supra* note 33, at 909 (“Once a court completes its assessment of the significance of the patented device, it should consider in addition the importance of the advance represented in the accused device.”).

C. Remedy

After duration and breadth, the third attribute of patent strength is remedy. Acting within the scope of a patent without a license from the owner infringes on it, and the law usually provides a remedy. For past infringements, the owner can sue to recover damages. In U.S. law, the usual basis for damages is compensation, which ideally puts the *victim* in the same position as if the infringement had not occurred.⁸⁷ Infringement would not have occurred if the user had bought a license. Therefore, compensation often equals the price of a license.

Higher damages than compensation (“super-compensatory”) are sometimes available. Treble damages—three times compensatory damages—is an optional remedy in U.S. patent law for *willful* infringement of patents.⁸⁸ Another super-compensatory remedy is disgorgement, which ideally puts the injurer in the same position as if the infringement had not occurred. Disgorgement is an unusual remedy, seldom available unless infringement involves a breach of a fiduciary duty or unjust enrichment.⁸⁹ Treble damages and disgorgement are two benchmarks for super-compensatory damages.

For future infringements, the owner traditionally has a choice of remedies. First, the owner can wait for infringement and recover damages after the fact. However, waiting is often impractical. Second, instead of waiting, the owner can enjoin future infringement before it occurs.⁹⁰ When facing the threat of an

⁸⁷ Compensatory damages differ in strength according to their basis of computation. For firms, compensation ideally equals the profits lost from infringement, (e.g., the decrease in the patent owner’s profits caused by the infringement). The owner’s profits equal the difference between revenues from selling the innovation and the cost of producing and developing it. Sometimes, the cost of the invention’s production and development is easier to document and prove than revenues foregone by sales that did not occur because of the infringement. In that case, the victim of infringement may ask for compensation of costs, not compensation of lost profits.

⁸⁸ 35 U.S.C. § 284 (2018) (in infringement cases, courts “may increase the damages up to three times the amount found or assessed”).

⁸⁹ See Amanda Frye, “*Inextricably Commingled*”: A Restitution Perspective in Patent Remedies, 26 HARV. J.L. & TECH. 669, 670 (2013) (“Since 1964, courts have evaluated patent remedies squarely from the torts perspective, embracing compensatory damages over disgorgement.”); Mark A. Lemley, *Distinguishing Lost Profits from Reasonable Royalties*, 51 WM. & MARY L. REV. 655, 655 (2009). Cf. Hacothen, *supra* note 14 (suggesting imposing disgorgement for pharmaceutical evergreening).

⁹⁰ In addition to patent law, U.S. firms have an additional remedy against infringing imports. The U.S. firm can bring a case before the U.S. International Trade Commission asking it to ban the import of all infringing goods. This powerful remedy against foreign infringement allegedly leads to abuse by firms

injunction, an infringer has no real choice but to desist from the infringing activity or pay the owner's asking fee for a license.

Until recently, injunctions were the chief remedy of future infringement in most patent cases.⁹¹ A seminal Supreme Court decision in the case of *eBay*, however, substitutes another remedy for injunction in some circumstances.⁹² The substitute remedy denies the owner an injunction against the infringer. Instead, the owner can collect a reasonable license fee, set by the court, for future use.⁹³ Thus, the user has a choice between *taking* a license at the court's price or *buying* a license at the owner's price. In Part V, we explore how we think the *eBay* case should be interpreted.

Now we will order remedies by strength. "No remedy" refers to an innovation that the law does not protect against unauthorized use. As depicted in Figure 9, "no remedy" is the weakest remedy. Moving to the right in Figure 9, "damages" refers to the user's choice to take the license at a court-set price (damages), buy a license at the owner's asking price, or not use it. If he decides to use the invention, he will usually choose the cheaper between take and buy. Next, an injunction against taking the invention forces the user to buy it at the owner's asking price, or not to use it. Consequently, an injunction is a stronger remedy than "buy or take." Finally, treble damage ideally equals three times compensation, which is usually higher than the owner's asking price. Thus, Figure 9 depicts the "treble damages" as stronger than "injunction."

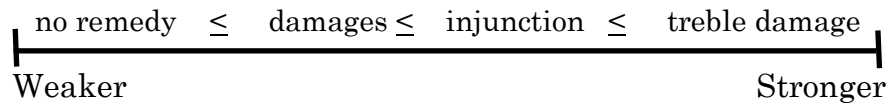
shielding themselves from foreign competition. See K. William Watson, *Still a Protectionist Trade Remedy: The Case for Repealing Section 337*, CATO INST. (2012), <https://object.cato.org/sites/cato.org/files/pubs/pdf/PA708.pdf>.

⁹¹ 35 U.S.C. § 261 (2018) ("[P]atents shall have the attributes of personal property."); Christopher B. Seaman, *Permanent Injunctions in Patent Litigation after eBay: An Empirical Study*, 101 IOWA L. R. 1949, 1960 (2016) ("Prior to *eBay*, courts routinely characterized patents as conferring a property right on their owners. In turn, the right to exclude has been widely viewed as the "hallmark of a protected property interest." (quoting *Fla. Prepaid Postsecondary Educ. Expense Bd. v. Coll. Sav. Bank*, 527 U.S. 627, 643 (1999)); *Dawson Chem. Co. v. Rohm & Haas Co.*, 448 U.S. 176, 215 (1980) (noting "the long-settled view that the essence of a patent grant is the right to exclude").

⁹² *eBay Inc. v. MercExchange, L.L.C.*, 547 U.S. 388 (2006) (A landmark decision in which the Supreme Court unanimously determined that an injunctive remedy should not be automatically issued based on a finding of patent infringement.).

⁹³ Cf. Guido Calabresi & A. Douglas Melamed, *Property Rules, Liability Rules, and Inalienability: One View of the Cathedral*, 85 HARV. L. REV. 1089 (1972) (arguing an injunctive regime can be viewed as a "property rule," whereas a damage regime can be viewed as a "liability rule.").

Figure 9. Strength of Remedies



Of these legal remedies, which one is best? According to the separation principle, infringing to innovate should usually have no remedy or a damage remedy, as depicted on the far left of Figure 9. At the same time, patent owners should have a strong remedy against infringement for production and consumption. How strong? One might assume that to maximize growth, patents should protect against consumption and production with the strongest remedy, shown on the far right in Figure 9. Professor Ted Baker, for example, has offered that a punitive damages remedy would protect patents for important inventions against certain infringements.⁹⁴

However, a patent owner's most preferred remedy for infringement is not necessarily the strongest remedy. To see why, consider how consumers and producers would respond to treble damages, which sharply sanction infringing. To avoid the risk of liability, potential patent users might avoid activities that possibly infringe, thus reducing demand for the innovation. Lower demand for the innovation reduces its owner's profits. Thus, the patent owner may prefer weaker damages to increase demand.

Instead of a punitive remedy, a better default remedy against uses in consumption and production is an injunction. A patent protected by an injunctive remedy gives monopoly power to the patent owner and allows him to settle his licensing terms in a way that maximizes his profits, including by stipulating a remedy.

A patent owner will often prefer stipulating his remedy as compensatory damages. Here's why: assume that the patent owner's license can stipulates damages for infringing. If licensed and unlicensed users have similar demands, the patent owner will maximize profits by setting damages roughly equal to the price of a license (in technical terms, the owner will not discriminate between licensed and unlicensed users).⁹⁵ Thus, the damages for infringement that maximize the patent owner's profits roughly equal the license fee, which roughly equals compensatory damages.

⁹⁴ Ted Baker, *Pioneers in Technology: A Proposed System for Classifying and Rewarding Extraordinary Inventions*, 45 ARIZ. L. REV. 445, 446 (2003) ("the proposed method would increase damages for the infringement of certain meritorious patents").

⁹⁵ A monopolist charges the same price if two groups of people have the same elasticity of demand and, if feasible, a different price when they have different elasticities of demand.

We have explained why the patent owner may prefer damages for infringement roughly equal to the license fee. Another consideration might cause the patent owner to prefer somewhat higher damages: The transaction costs of litigating are higher than those for buying. Remedies should ideally channel users into licensing rather than litigating. Setting damages above the license fee gives the user incentive to license rather than infringe and compensate. Therefore, the best damages usually equal the license fee plus a penalty for failing to license.

We explained that injunctions allow patent owners to stipulate their preferable remedy to maximize their profits. The next question then is, should the courts defer to the patent owner's stipulated remedy or set it aside in favor of a legal remedy?⁹⁶ If the courts are willing to enforce the stipulated remedy, then the legal remedy is a *default* rule. Conversely, if the courts are unwilling to enforce the stipulated remedy, then the legal remedy is a *mandatory* rule.⁹⁷

The court's deference or disregard for a stipulated remedy should depend on whether the remedy that is best for the patent owner is also best for society. According to the overtaking principle, maximizing innovation will increase social welfare. To fulfill its constitutional purpose and to increase social welfare, the law should defer to stipulations when they maximize growth. Otherwise, the law should mandate the remedy.

According to the separation principle, the law should distinguish consuming and producing from innovating. The law should defer to stipulated remedies against infringing on the patent by consuming and producing. Conversely, the law should mandate a weak remedy against infringing on the patent by innovating.⁹⁸ In sum, the courts should assume that remedies stipulated in patent

⁹⁶ The enforceability of stipulated remedies is a large topic that exceeds the scope of this Article. We note that courts traditionally enforce stipulations of compensation for breach of contract, and courts traditionally do not enforce stipulations of punishment or damages for tortious accidents.

⁹⁷ If the legal remedy is the one preferred by the patent owner, then the patent owner has no incentive to stipulate an alternative in a license. Conversely, if the legal remedy is not the one preferred by the patent owner, then he has an incentive to stipulate an alternative. By providing the legal remedy that the parties prefer, the law makes the stipulation of a remedy unnecessary, which reduces the negotiation costs of patent licensing. See ROBERT D. COOTER & THOMAS S. ULEN, *LAW AND ECONOMICS* 294–99 (6th ed. 2012), available at <https://scholarship.law.berkeley.edu/cgi/viewcontent.cgi?article=1001&context=books>. Alternatively, by providing the legal remedies that the parties dislike, the law incentivizes the parties to stipulate a remedy. Ian Ayers & Robert Gertner, *Strategic Contractual Inefficiency and the Optimal Choice of Legal Rules*, 101 *YALE L.J.* 729 (1992) (exploring the role of “penalty default” rules).

⁹⁸ See *infra* Part V.

licenses for consuming and producing are socially optimal, whereas remedies stipulated in patent licenses for innovating are socially excessive.⁹⁹

V. DOCTRINE AND POLICY

We explained that patent strength should be strong against production and consumption and weak against innovation. This section translates economic reasoning into legal practice. In Section V.A, we explore how legal doctrines that govern patents' scope and remedy could be utilized to promote separation.¹⁰⁰ Then, in Part V.B, we discuss exceptions to separation that involve holdout and uncertain rights. In these special circumstances, patent policy might also benefit from weakening patent rights against production and consumption, not just innovation.

Courts and commenters often confuse these two fundamentally different justifications for weakening patent rights. The former lies at the heart of patent policy but is nevertheless underutilized. The latter deals with generic market failures but is overutilized. We criticize this unfortunate trend.

A. *Weakening Patents Against Innovation*

Adhering to the separation principle, legal doctrines and policies should provide strong patent protection against consuming or producing and weaker patent protection against innovating. Legal doctrines can achieve this goal by empowering the courts to retain the full patent strength against consuming or producing and to weaken the patent strength (breadth and remedy) against innovating. We divide the discussion into three major categories of innovative uses that justify weakening patent strength: idea piggybacking, experimenting, and improving. For each category, we emphasize places where patent policy already achieves separation and places where legal policy should be further tailored to achieve this purpose.

⁹⁹ Patents sometimes confer a legal monopoly. Consequently, a patent license is a contract with a monopolist. Monopolists often offer form contracts with nonnegotiable terms. A contract with nonnegotiable terms is called a "contract of adhesion" in legal literature. In a contract of adhesion, the monopolist often gains the most by setting an *inefficiently* high price and efficient non-price terms. In some circumstances, however, the monopolist gains the most by setting inefficient non-price terms. For an economic analysis of contracts of adhesion, see COOTER & ULEN, *supra* note 97, at 363–67.

¹⁰⁰ We will expand here only on few policy levers that we consider to be highly fundamental for the separation function. For an extensive analysis of policy levers in patent policy, see BURK & LEMLEY, *supra* note 9.

i. Idea Piggybacking

Innovation sometimes motivates and spurs further innovation. Reading patent documents may expose prospective innovators to novel and innovative solutions for technological problems, which in turn may spark new ideas and seed future innovation.¹⁰¹ This process, in which ideas disclosed by a previous innovator serve as input in a mental progress done by a later innovator, was named by Professor Collins “idea-only cumulative innovation cases.”¹⁰² Professor Collins explained these cases by providing an example:

[A]n earlier inventor may patent a molecule that is useful for treating a disease. The patent specification discloses the structure of the molecule. After the patent has been filed, someone else may learn the knowledge of the protein structure disclosed in the patent specification, appropriate an idea qua idea from the earlier patent, and have a “Eureka!” moment. He may realize that if a molecule with the molecular structure revealed in the patent has a particular biological activity, then perhaps the metabolic pathway in a cell must include a particular step in order for the molecule to have that activity. The later actor may go on to discover a previously unknown metabolic pathway.¹⁰³

While the “Eureka” moment of the junior party was evidently facilitated by the work of the senior party, it is nevertheless clear that the latter is not protected in any way against such a use of his innovative input.¹⁰⁴ Indeed, patent policy allows for free piggybacking on ideas that were disclosed by earlier inventors. Thus, as Professor Collins stated, it will be a lawful business to communicate disclosed information from published patent documents for profit without fearing patent infringement.¹⁰⁵

Idea piggybacking is allowed because patent protection does not extend to ideas. Instead, patents only provide a right to exclude from

¹⁰¹ See Benjamin N. Roin, *The Disclosure Function of the Patent System (Or Lack Thereof)*, 118 HARV. L. REV. 2007, 2010 (2005) (“[T]he patent may ‘give ideas’ to engineers in other industries who would otherwise never hear about the technology, and thus allows everyone the opportunity of discovering new ways of using it.”).

¹⁰² Collins, *supra* note 33, at 1249.

¹⁰³ *Id.* at 1250.

¹⁰⁴ *Id.*

¹⁰⁵ *Id.* at 1249 n.117; see also *Teletronics Pacing Sys., Inc. v. Ventritex, Inc.*, 982 F.2d 1520, 1523 (Fed. Cir. 1992) (holding that the dissemination of data about a device falling within a patent’s claims is not an infringing activity).

a set of things (and processes) that embody innovative ideas.¹⁰⁶ The economic justification for the lack of patent protection to ideas is grounded in the separation principle. Using knowledge to fuel creativity and spark ingenuity is a purely innovative activity performed by innovators.¹⁰⁷ Thus, patents are aptly granted zero protection against ideas per se.

ii. *Experimenting*

Once an idea to innovate has formed in an innovator's mind, the next step requires practicing that idea. As depicted in Figure 1, the first phase of practicing innovation is "development," which often includes experimenting with patented goods. Experimenting to improve a drug is attempted innovation. Consequently, the separation principle favoring innovation applies. An ideal patent policy would weaken patent rights in favor of experimenting. Thus, the separation principle supports an exception for infringing to experiment. In notation, recall our example in Figure 8 of senior party A who owns the patented invention α and junior party B who uses α to make the allegedly infringing invention β . Patent policy should weaken protection of A's rights against B's experimental use of α . Removing experimentation from liability encourages follow-on innovation by B.¹⁰⁸

Historically, under the common law, patent policy has permitted wide, unlicensed use of patented goods for noncommercial experiments by universities and other nonprofit organizations. The origin of the common law's experimental use defense is widely attributed to several court decisions by Justice Story in the early nineteenth century.¹⁰⁹ In Story's view, "it could never have been the intention of the legislature to punish a man, who constructed a [patented] machine merely for philosophical

¹⁰⁶ Collins, *supra* note 33, at 1250; Rubber-Tip Pencil Company v. Howard, 87 U.S. (20 Wall.) 498, 507 (1874) ("An idea of itself is not patentable, but a new device by which it may be made practically useful is.").

¹⁰⁷ Indeed, fertilizing innovation through disclosure is one of the basic premises on which patent policy is built. See Roin, *supra* note 101, at 2009; Brett M. Frischmann & Mark A. Lemley, *Spillovers*, 107 COLUM. L. REV. 257 (2007) (explaining why positive externalities are good for fertilizing innovation); see also Kewanee Oil Co. v. Bicron Corp., 416 U.S. 470, 481 (1974) (explaining that the disclosure of patent applications adds to the public's "general store of knowledge" and "is assumed [to] stimulate ideas and the eventual development of further significant advances in the art.").

¹⁰⁸ The effect of a research exemption also depends on other patent rights, specifically whether an improvement will infringe on the prior patent ("blocking patents" scenario), as we explain in the next section. See also SUZANNE SCOTCHMER, *INNOVATION AND INCENTIVES*, Ch. 5, 128–157 (2004).

¹⁰⁹ Whittemore v. Cutter, 29 F. Cas. 1120 (C.C.D. Mass. 1813) (No. 17,600); Sawin v. Guild, 21 F. Cas. 554 (C.C.D. Mass. 1813) (No. 12,391).

experiments [today's 'science'],¹¹⁰ or for the purpose of ascertaining the sufficiency of the machine to produce its described effects.”¹¹¹ Story favored weak patent protection against experimentation, as prescribed by the separation principle.¹¹²

Unfortunately, courts have sharply curtailed the experimental use defense over the years.¹¹³ In 1984, the Federal Circuit ruled in *Roche v. Bolar*¹¹⁴ that the experimental use defense did not protect a generic drug manufacturer experimenting to obtain FDA approval.¹¹⁵ In 2000, the Federal Circuit further curtailed the defense in *Embrex v. Service Engineering Corp.*,¹¹⁶ finding that experimenting to design around a patent infringes. Finally, in 2002 the Federal Circuit “seal[ed] the coffin of the experimental use

¹¹⁰ See *Integra Lifesciences v. Merck*, 331 F.3d 860, 875 n.8 (Fed. Cir. 2005) (Newman, J. Dissenting) (“By ‘philosophical’ experiments, Justice Story was referring to ‘natural philosophy,’ the term then used for what we today call ‘science’.”).

¹¹¹ *Whittemore*, 29 F. Cas., at 1121.

¹¹² For a comprehensive review of the history of the experimental use privilege, see TED HAGELIN, NYSTAR RESEARCH REPORT, THE EXPERIMENTAL USE EXEMPTIONS TO PATENT INFRINGEMENT (2005), <http://nysstlc.syr.edu/wp-content/uploads/2014/09/NYSTAR-The-Experimental-Use-Exemptions-to-Patent-Infringement-Fall-2005.pdf>.

¹¹³ Shortly after Story's early decisions, the experimental use defense was narrowed in *Poppenhusen v. Falke*, 19 F. Cas. 1048 (C.C.N.Y. 1861) (No. 11279) (the experimental use defense will accommodate only uses for “the sole purpose of gratifying a philosophical taste, or curiosity, or for mere amusement”); then the interpretation was expended once again beyond “mere amusement,” in few subsequent decisions. See *Ruth v. Stearns-Roger Mfg. Co.*, 13 F. Supp. 697 (D. Colo. 1935) (a university's use of an infringing machine parts solely in a laboratory and for the sole purpose of experimentation is not infringing); *Akro Agate Co. v. Master Marble Co.*, 18 F. Supp. 305, 333 (1937) (experimenting with a patented machine before commercial production is not infringing); *Dugan v. Lear Avia, Inc.*, 55 F. Supp. 223, 229 (1944) (construction of a single patented device for experimentation without ever manufacturing devices for sale is not infringing). The defense was narrowed again soon thereafter in *Pitcairn v. United States*, 546 F.2d 1106 (1977), in which the government used patented helicopter rotors and controls for testing and evaluation of such factors as lift ability, vibration, flight speed, and range. The court held that testing and evaluation were “intended uses of the infringing aircraft . . . and are in keeping with the legitimate business of the using agency,” and thus infringing. *Id.* at 1125-26. Following this decision, experimentation that is related to the experimenter's legitimate business is not exempted from liability regardless of whether the use was commercial or noncommercial. See David L. Parker, *Symposium, Patent Infringement Exemptions for Life Science Research*, 16 Hous. J. Int'l L. 615, 631 (1994).

¹¹⁴ 733 F.2d 858 (Fed. Cir. 1984).

¹¹⁵ The court reached this decision knowing that establishing bioequivalency in generics (as was attempted in this case) was a mandated regulatory requirement and a prerequisite for entering the drug market. *Id.* at 860. The court further acknowledged that the result would be a de facto extension of the patent term in favor of the brand name companies since generics will have to postpone the regulatory process until the patent expiration. *Id.* at 864.

¹¹⁶ 216 F.3d 1343 (2000).

exception”¹¹⁷ in *Madey v. Duke*¹¹⁸ by excluding any use with the “slightest commercial implication” or “in keeping with the legitimate business of the alleged infringer.”¹¹⁹

As the judiciary pruned the experimental use defense, the legislature partially restored it for pharmaceuticals. In 1984, the Hatch-Waxman Act reversed the holding of *Bolar*.¹²⁰ The statute permits generic drug companies to experiment with brand-name drugs without a license before the patent expires.¹²¹ Consequently, generics can get regulatory approval sooner and begin competing with brand-name drugs earlier.

According to the separation principle, the research exemptions should be restored for innovating, as with universities. Innovating, however, is not imitating. When applying the separation principle, imitating is a form of producing. Imitation increases production without creating anything new. In the case of patented drugs, generic imitation increases production and reduces the drug’s price, which reduces profits from the original patent. Reduced profits for innovation imply less of it, as depicted in Figure 3. Thus, the Hatch-Waxman Act partly restored the experimentation exemption to encourage generic imitations and drive down patented drug prices. Slowing down innovation was its likely effect.

According to the overtaking principle, the social loss from slower innovation overtakes the social gain from lower prices (Figure 6.). Consequently, the experimental use privilege should not

¹¹⁷ Petition for Writ of Certiorari at 14, *Duke Univ. v. Madey*, 539 U.S. 958 (2003) (No. 02-1007) (denying cert.).

¹¹⁸ *Madey v. Duke Univ.*, 307 F.3d 1351 (Fed. Cir. 2002).

¹¹⁹ *Id.* at 1362. According to the majority opinion, even though Duke University clearly used the patented technology for scientific research purposes, such use was not protected as it could not be separated from the university’s “business objectives,” which include educating students and faculty, enhancing the status of the institution, and luring lucrative research grants, students, and faculty. *Id.*

¹²⁰ Drug Price Competition and Patent Term Restoration Act (Hatch-Waxman Act), Pub. L. No. 98-417, 98 Stat. 1585 (1984) (codified as amended in scattered sections of 15, 21, 28 and 35 U.S.C.). For a description of the act’s goal and misuse, see Hacothen, *supra* note 14.

¹²¹ In the Hatch-Waxman Act, Congress amended section 271(e)(1) of the Patent Act 35 U.S.C. to provide that “it shall not be an act of infringement to make, use, offer to sell, or sell . . . a patented invention . . . solely for uses reasonably related to the development and submission of information under a Federal law.” This so-called “Bolar exemption” was designed to hasten generic competition by enabling generic competitors to experiment with patented drugs for the purpose of regulatory compliance within the term of patent protection. *See Integra Lifesciences v. Merck*, 331 F.3d 860, 865 (Fed. Cir. 2003), *vacated*, 545 U.S. 193 (2005) (“The House Committee that initiated this provision expressly described the pre-market approval activity as ‘a limited amount of testing so that generic manufacturers can establish the bioequivalency of a generic substitute.’”), citing H.R. Rep. No. 857, at 8, *reprinted* in 1984 U.S.C.C.A.N. at 2692.

apply to experimenting to imitate. The Hatch-Waxman exception cannot be sustained under the overtaking principle as advanced in this paper (although it might be justified on administrative reasons).¹²²

Fortunately, the Supreme Court partially restored the research exemption for improving drugs, not just imitating them. Under *Merck v. Integra*,¹²³ competing innovators enjoy the privilege of using patented drugs without a license when trying to improve them.¹²⁴ Judge Newman's concurring opinion in the lower *Merck* decision is especially illuminating. Newman stated:

The subject matter of patents may be studied in order to understand it, or to improve upon it, or to find a new use for it, or to modify or "design around" it. Were such research subject to prohibition by the patentee the advancement of technology would stop, for the first patentee in the field could bar not only patent-protected competition, but all research that might lead to such competition, as well as barring improvement or challenge or avoidance of patented technology. Today's accelerated technological advance is based in large part on knowledge of the details of patented inventions and how they are made and used. Prohibition of research into such knowledge cannot be squared with the framework of the patent law.¹²⁵

Still, the statutory exception as interpreted in *Merck* is nevertheless limited in both its scope (it does not extend to basic scientific research)¹²⁶ and reach (it applies only to pharmaceuticals and related medical products). Outside that reach, the experimental

¹²² Indeed, the Hatch-Waxman Act upset patent principles in an attempt to restore the balance that was lost by introducing the complex regulation of prescription drugs by the FDA. See Hacothen, *supra* note 14.

¹²³ *Merck KGaA v. Integra Lifesciences I, Ltd.*, 545 U.S. 193 (2005).

¹²⁴ Under the Court's interpretation of 271(e)(1), the use of patented subject matter in (i) experimenting on drugs that are not ultimately the subject of an FDA submission and in (ii) obtaining research data that is not ultimately submitted to the FDA can both be exempted from infringement under certain conditions. *Id.* at 206.

¹²⁵ *Integra Lifesciences I, Ltd. v. Merck KGaA*, 331 F.3d 860, 875 (Fed. Cir. 2003), *vacated*, 545 U.S. 193 (2005).

¹²⁶ *Merck*, 545 U.S. at 205–06 (saying that "[b]asic scientific research on a particular compound, performed without the intent to develop a particular drug or a reasonable belief that the compound will cause the sort of physiological effect the researcher intends to induce" is outside the bounds of the statutory safe harbor).

use exception is all but eliminated under the rule of *Madey v. Duke*.¹²⁷

According to the separation principle, patents should be weak against experimentation to innovate, and the experimental use defense can be interpreted to achieve this result. Lawmakers should consider adopting an experimental use defense that applies to all innovations, not just pharmaceuticals.¹²⁸ In fact, outside the arena of pharmaceuticals, in industries characterized by a cumulative and incremental innovative process such as in the software industry, the experimental use defense should play an even more prominent role.

Professor Donald Chisum, for example, suggested that the experimental use defense should be interpreted to exclude from patent infringement tinkering with a patented algorithm during research.¹²⁹ Professors Mark Lemley and Julie Cohen also advanced a similar proposal.¹³⁰ Also, many European countries already have expansive experimental use privileges.¹³¹ In Germany, for example,

¹²⁷ See HAGELIN, *supra* note 112, at 13 (“The sweeping holding in *Madey v. Duke* would appear to preclude experimental use of patented subject matter by all non-profit research organizations, including federal laboratories, research foundations and research hospitals. . . it could conceivably be interpreted to preclude experimental use of patented subject matter even by isolated individuals if the use was pursuant to any specific objective”).

¹²⁸ Many scholars have offered suggestions along these lines. For a summary, see Peter S. Menell & Suzanne Scotchmer, *Intellectual Property*, in THE HANDBOOK OF LAW AND ECONOMICS 1473 (A. Mitchell Polinsky & Steven Shavell eds., 2007).

¹²⁹ See Donald S. Chisum, *The Patentability of Algorithms*, 47 U. PITT. L. REV. 959, 1017–18 (1986).

¹³⁰ See Mark A. Lemley & Julie E. Cohen, *Patent Scope and Innovation in the Software Industry*, 89 CALIF. L. REV. 1, 29 (2001) (arguing that the experimental use defense should be interpreted to allow reverse engineering of software).

¹³¹ See Rebecca S. Eisenberg, *Patent Swords and Shields*, 299 SCI. 1018, 1018 (2003); Janice M. Mueller, *The Evanescent Experimental Use Exemption from United States Patent Infringement Liability: Implications for University and Nonprofit Research and Development*, 56 BAYLOR L. REV. 917, 918 (2004). Proposals by the European Union Parliament clearly equate experimental use provisions with increased competitiveness. See, e.g., Natalie M. Derzko, *A Local and Comparative Analysis of the Experimental Use Exception—Is Harmonization Appropriate?*, 44 IDEA 1, 62 (2003); John A. Tessensohn, *Reversal of Fortune—Pharmaceutical Experimental Use and Patent Infringement in Japan*, 4 J. INTL. LEGAL STUD. 1, 38 (1998); Kevin Sandstrom, *How Much Do We Value Research and Development?: Broadening the Experimental Use Exemption to Patent Infringement in Light of Integra Lifesciences I, Ltd. v. Merck KGaA*, 331 F.3d 860 (Fed. Cir. 2003), 30 WM. MITCHELL. L. REV. 1059 (2004) (arguing that countries with high innovative capacities all have broad experimental use exemptions in their law, citing Germany, Japan, Italy, Switzerland, France, Spain, Sweden, Canada, and the United Kingdom). But see Kevin Iles, *A Comparative Analysis of the Impact of Experimental Use Exemptions in Patent Law on Incentives to Innovate*, 4 NW. J. TECH. & INTELL. PROP. 61, 81 (2005) (“[C]ontrary

scientific experimentation is permitted, regardless of the underlying commercial purpose.¹³²

The experimental use defense promotes separation by tailoring the *scope* of patents against experimentation that leads to further innovation.¹³³ Other doctrines such as preemption or patent misuse can also be tailored to achieve the same result.¹³⁴ An alternative approach would be to tailor patents' *remedy* against experimentation. Thus, Professor Rebecca Eisenberg offered that in some instances, experimentation would be infringing but should nevertheless be remedied by damages, not an injunction.¹³⁵

Finally, patent policy could also promote separation by shortening the *duration* of certain categories of inventions that are prone to experimentation such as software patents or research tools.¹³⁶ Different approaches may also complement one another. Professor Strandburg, for example, offered that patent duration for research tools should be shorter initially but then complemented by an additional period of protection limited to a damage remedy.¹³⁷

How far should policy makers go when weakening patents against experimentation? A full analysis requires confronting the fertility principle. For example, even though research tools are used exclusively for innovation (which justifies weaker protection, according to the separation principle), such inventions are more fertile than many of their uses, which justifies tailored protection

to predictions in the literature, a narrow experimental use exemption might not correlate with reduced R&D.”); Norman Siebrasse & Keith Culver, *The Experimental Use Defense to Patent Infringement: A Comparative Assessment*, 56 U. TORONTO L. J. 333, 335 (2006) (arguing that overall “the European experience with the experimental use defense provides a cautionary tale rather than an example to be followed”).

¹³² *Klinische Versuche I* [Clinical Trials I], Bundesgerichtshof [BGH] [Federal Court of Justice] Oct. 30, 1997, R.P.C. 623, 639 (Ger.); *see also* Iles, *supra* note 131, at 77–79 (exploring the German experimental use provision); *see also* András Kupecz et al., *Safe Harbors in Europe: An Update on the Research and Bolar Exemptions to Patent Infringement*, 33 NATURE BIOTECHNOLOGY 710 (2015), <https://www.nature.com/articles/nbt.3273>.

¹³³ *See supra* note 85 and accompanying text (classing exceptions to infringement liability as part of the patent’s scope).

¹³⁴ *Cf.* Lemley & Cohen, *supra* note 130, at 30–36 (arguing the doctrines of preemption and patent misuse can also be adjusted to advance the same policy goals pursued by the experimental use defense).

¹³⁵ *See* Rebecca S. Eisenberg, *Patents and the Progress of Science: Exclusive Rights and Experimental Use*, 56 U. CHI. L. REV. 1017, 1077 (1989) (“[I]n some cases it may be appropriate for a court to require the researcher to pay a reasonable royalty to the patent holder after the fact in order to be sure that the patent holder is adequately compensated for the use of the patented invention.”).

¹³⁶ *See* notes 46–48 and accompanying text.

¹³⁷ *See* Katherine J. Strandburg, *What Does the Public Get? Experimental Use and the Patent Bargain*, 1 WISC. L. REV. 82, 142 (2004).

over complete exclusion.¹³⁸ The optimal patent strength against experimentation would be impacted by factors such as: (1) whether the experimentation leads to an improvement¹³⁹ and if so, (2) the fertility of that improvement. We explore the latter issue briefly in the next section.¹⁴⁰

Investigating these issues in depth goes beyond the scope of this Article because we focus on separation, not fertility. Nevertheless, wherever the optimal balance lies, it is safe to say that in contrast to the separation mandate, current legal policy fails to weaken patent strength against experimentation.¹⁴¹

iii. Improving

An improvement adds to an old invention. Both the original inventor and the improver usually contribute to an improvement. In principle, patent law could give exclusive ownership rights over improvements to the original inventor, not to the improver.¹⁴² If so, each user, including the improver, would have to negotiate a license with the original patentee alone. To illustrate, assume that A invents α and patents it, and B subsequently invents an improvement β and attempts to patent it. The law can recognize A's patent over α as encompassing all ownership rights over β and deny a separate patent for β . In that case, B is unable to get a patent over β or use it without A's authorization. A has sole dominion over the original invention

¹³⁸ Cf. Eisenberg, *supra* note 135, at 1078 (suggesting that experimentation should not be excused from infringement liability when a patented invention has “a primary or significant market among research users”).

¹³⁹ Indeed, if the experimentation process does not lead to a marketable improvement, there is little reason to deny full exemption from liability for the experimentation. Cf. Eisenberg, *supra* note 135, at 1078 (suggesting that experimentation done “to check the adequacy of the specification and the validity of the patent holder’s claims about the invention should be exempt from infringement liability”).

¹⁴⁰ See *infra* note 157 and accompanying text.

¹⁴¹ If not through court discursion, a legislative action might be needed to weaken patent strength against experimentation. Cf. Lemley & Cohen, *supra* note 130 (urging a legislative change to accommodate reverse-engineering of software); Peter S. Menell & Jeffrey A. Lefstin, *Reforming Patent Eligibility: Supplementary Statement* (U.C. Hastings Research Paper No. 336, 2019), <https://ssrn.com/abstract=3399499> or <http://dx.doi.org/10.2139/ssrn.3399499> (urging the Senate Judiciary Subcommittee to consider addressing the issue of experimental use).

¹⁴² This view was advanced by Professor Edmund Kitch, who argued that patents should assume a “prospect” function. According to Kitch, in order to provide optimal incentive to create the pioneer innovation, the pioneer inventor should be able to capture all of the social benefits derived from the invention, including those that flow from improvements. Edmund W. Kitch, *The Nature and Function of the Patent System*, 20 J. L. & ECON. 265 (1977).

α as well as over any improvement such as β (copyright takes a similar approach.)¹⁴³

Alternatively, the symmetrically opposite approach gives exclusive ownership rights over improvements to the improver, not to the original inventor. Each user, including the original inventor, would have to negotiate a license with the improver alone. To illustrate, as before, assume that A invents α and patents it, and B subsequently invents an improvement β and attempts to patent it. The law can recognize B's patent as encompassing all ownership rights over β and deny that A's patent over α covers β . In that case, A is unable to use β without B's authorization. B has sole dominion over the improvement β .

This is the usual result when an initial invention remotely influences a subsequent invention. Instead of an "improvement," the new invention is legally "independent." As Professor Collins put it, "[t]he earlier invention of a particular type of blast furnace may facilitate the later invention of a new type of metal alloy, but the metal alloy is not an improvement on the furnace."¹⁴⁴

Sometimes, however, the law takes an intermediate position between exclusive ownership by the original inventor and the improver. Instead of exclusive ownership, the law sometimes gives them joint ownership. The improver receives a patent without the original inventor's consent.¹⁴⁵ The original innovator A and the improver B jointly own the improvement. Each one needs a license

¹⁴³ When a book is adapted into a film without authorization or an unauthorized sequel to a movie is created, the newly added features will not be protected by copyright law, no matter how creative they are (unless the use is qualified as fair use). See 17 U.S.C. § 103(a) (2018) (providing that "[t]he subject matter of copyright as specified by section 102 includes compilations and derivative works, but protection for a work employing preexisting material in which copyright subsists does not extend to any part of the work in which such material has been used unlawfully"); Lemley, *supra* note 25, at 1021. If, however, part of the derivative work is independent and does not incorporate any infringing materials, these parts may be entitled to copyright protection. See *id.* at 1022.

¹⁴⁴ See Collins, *supra* note 33, at 1246.

¹⁴⁵ To understand how an unauthorized improvement can happen in practice, consider this historic example: The Thompson brothers obtained a patent on shock absorbers for motor vehicles in 1913. The Thompson patent enjoyed significant commercial success initially but was soon discovered to be ineffective due to wear and tear on the shock absorbers that resulted from friction. In 1918, William Storrie obtained a patent on an improved shock absorber that remedied the problem. The Thompson brothers actively challenged Storrie's patent for his improvement, arguing that they were its rightful owners. In spite of opposition by the Thompsons, the Patent Office issued the patent to Storrie. *Temco Elec. Motor Co. v. Apco Mfg. Co.*, 275 U.S. 319, 322-26 (1928) (describing the interference proceedings between the Thompsons and Storrie).

from the other to practice the improvement.¹⁴⁶ The literature calls this situation “blocking patents.” Every patent blocks unauthorized use, but “blocking patents” refers to overlapping patents in which each one blocks the other’s use.¹⁴⁷

In our notation, if pioneer A receives a patent for α that also covers β , and improver B subsequently receives a patent for β , each one can exclude the other from practicing β . To make this more concrete, imagine that A had a pioneering patent over a new substance α , and that B obtained a patent over a better process β for making α . In such a case, making α by process β will require the consent of both A and B.¹⁴⁸ Commercialization requires cooperation.

In principle, the parties can cooperate with each other, regardless of whether the patent rights are exclusive or joint. In either case, the Coase Theorem asserts that the original inventor and

¹⁴⁶ Indeed, improvement patents are an integral part of the patent system. *See In re Hogan*, 559 F.2d 595, 606 (C.C.P.A. 1977) (“[E]ncouragement of improvements on prior inventions is a major contribution of the patent system and the vast majority of patents are issued on improvements.”); John F. Duffy, *Inventing Invention: A Case Study of Legal Innovation*, 86 TEX. L. REV. 1, 32 (2007) (noting that starting in the late eighteenth century, British courts rejected “hostility to improvement patents”); John M. Golden, “Patent Trolls” and Patent Remedies, 85 TEX. L. REV. 2111, 2117 (2007) (showing that improvement patents were recognized for centuries).

¹⁴⁷ Patents confer on their owners the negative right to exclude others, usually through the threat of an injunctive relief. *Cf.* Richard Epstein, *The Disintegration of Intellectual Property? A Classical Liberal Response to a Premature Obituary*, 62 STAN. L. REV. 455, 493 (2010). Patents, however, do not confer on their owners the positive right to practice the invention. The legal term “practice” encompasses making, manufacturing, and reproducing. *See* Richard H. Shear & Thomas E. Kelley, *A Researcher’s Guide to Patents*, 132 PLANT PHYSIOLOGY 1127 (2003) (“[P]atents are negative rights, not positive privileges. A patent provides its owner with the right to prevent the manufacture, sale, use, importation, or offer for sale of the patented product, process, or composition. A patent is not a positive right that enables its owner to do anything that he or she wants.”). Thus, the fact the improver had a patent over his improvement does not necessarily mean that he can also practice the improvement unconditionally.

¹⁴⁸ Another example of an improvement scenario is one in which A obtained a pioneer patent over a class or a *genus* of substances (α), and then B obtained an improvement patent over a superior single substance or *species* within that class (β). In such a case, the improvement is β (i.e. the superior substance), but it will nevertheless be impossible to practice the improvement without running afoul of the pioneer patent (α). Therefore, similar to the example we used above, the consent of both A and B is needed. *See* Ariad Pharm., Inc. v. Eli Lilly & Co., 598 F.3d 1336 (Fed. Cir. 2010). For criticism on this type of blocking situation, *see* Jason Rantanen, *The Doctrinal Structure of Patent Law’s Enablement Requirement*, 69 VAND. L. REV. 1679, 1712 (2016) (“[W]hen a genus is viewed as being made up of identifiable constituents, those individual ‘species’ will tend to present individual targets, such that each must be enabled.”).

improver will cooperate to create value unless transaction costs inhibit negotiations.

To illustrate cooperation, consider unblocking by contract. Assume that A makes the original invention α , B makes the improvement β , and they have overlapping patents to β . To practice the improvement, they need a cross-license. Both A and B can gain by cross-licensing and then dividing the surplus. Bargaining for a license usually succeeds as long as the process of negotiating is not too costly. Applied to blocking patents, the Coase Theorem asserts that whenever unblocking a patent creates value, contracts will unblock it unless transaction costs inhibit negotiations.

We have discussed cooperation under joint ownership of an invention. Cooperation under exclusive ownership by the original inventor is also possible but is difficult to accomplish by private contract. In Coase's language, the transaction costs are excessive. Here is why: Developing an improvement requires up-front investment. Once developed, imitating the improvement is relatively cheap. Exclusive ownership of the improvement by the original inventor gives the improver no shield against the original inventor's imitating the improvement. Therefore, a potential improver is reluctant to improve for fear of imitation by the original inventor.¹⁴⁹

The law solves the problem by giving a patent to the improver. Patentability encourages improvements by precluding

¹⁴⁹ A prospective improver would not have a patent right to secure the information regarding his improvement from the original inventor during negotiation. This would make it very costly for him to communicate the potential commercial value of the improvement without surrendering the information on how to make the improvement. Once this information is disclosed, the pioneer has little reason to pay for it. Consequently, the negotiations between pioneer and improver would probably fail. This is the "Arrow problem" of information applied to improvements. See Kenneth J. Arrow, *Economic Welfare and the Allocation of Resources for Invention*, in *THE RATE AND DIRECTION OF INVENTIVE ACTIVITY: ECONOMIC AND SOCIAL FACTORS* 609, 615 (National Bureau of Economic Research ed., 1962); Lemley, *supra* note 33, at 1051 ("The patent owner has no good way of evaluating potential improvers and choosing the most efficient one (or ones). It would be far more efficient to find a way for improvers to invent first, and then bargain with the original patent owner over the allocation of rights to the improved invention. The original inventor would only have to deal with a small subset of all potential improvers—those that were successful."); *IP Bargaining Breakdown*, *supra* note 33, at 80–81. Note that experimentation to improve makes even less sense if it is not exempted from liability—that is, in the absence of an experimental use defense.

imitation. Thus, blocking patents stimulate improvements just as patents stimulate original innovations.¹⁵⁰

Besides making cooperation more likely, blocking patents affect the distribution of the resulting surplus. Compared to exclusive ownership by the original inventor, a blocking patent gives the improver more leverage in negotiations with the original patentee.¹⁵¹ By lowering the fee that an improver must pay for a license, blocking patents redistribute wealth from the original inventor to the improver.¹⁵² However, when negotiating a license,

¹⁵⁰ See *IP Bargaining Breakdown*, *supra* note 33, at 95 n.80 (“By enhancing the bargaining position of improvers, improvement patents make cross-licenses more likely. They represent an intriguing and successful tailoring of property rights to a difficult private bargaining situation.”).

¹⁵¹ Other than removing the threat of imitation, a blocking patents situation enhances the bargaining position of the improver against the original patentee in other respects as well. First, the original patentee’s threat to sue for infringement is lower in these cases because the patentable nonobvious improvement might indicate to the court that there was no infringement. *Cf. Zygo Corp. v. Wyko Corp.*, 79 F.3d 1563 (Fed. Cir. 1996) (suggesting that an improvement’s patentability is evidence of noninfringement). Second, even if the improver is indeed an infringer, the original patentee’s threat of damages is also lower. The patented improvement will lower the amount of damages that the original patentee will be able to recover for past infringement since the original patentee will not be able to argue that the sales of his original invention (α) might have replaced all of the sales of the improvement (β). This is because some of these sales are arguably attributed to the value added by the improvement (β). *Merges & Nelson*, *supra* note 33, at 862. Third, if the improvement is “radical,” the court could liberate the improver from liability altogether, using the reverse doctrine of equivalence. See *Rent Control*, *supra* note 33; *IP Bargaining Breakdown*, *supra* note 33. We will discuss this option further below. See *infra* note 153 and accompanying text.

¹⁵² While the blocking patent doctrine weakens the rights of the original patentee relative to the improver by eliminating the risk of imitation, there is a significant practical limitation to that privilege. The costs for coming up with the improvement are sunk by the time the improver can negotiate with the original patentee for a license. The original patentee can leverage the improver’s sunk costs to extract a better bargain. See *Siebrasse & Culver*, *supra* note 131, at 351 (“[T]he improver faces a major disadvantage in *ex-post* bargaining that counters the patent-in-hand advantage, namely the sunk costs incurred in developing the improvement. Because the costs are sunk the time of *ex-post* negotiations, if the improver walks away without a deal it will suffer a loss equal to the entire development costs. This means that even if the negotiations are successful, the improver is not guaranteed net profit, as it is better off accepting a deal that gives it something rather than nothing, even if its share does not allow it to recover its full development costs.”). Thus, while the blocking patent doctrine always improves the standing of the improver by allowing him to opt for *ex post* licensing rather than *ex ante* licensing, the improver is likely to exercise this privilege only if the improver’s leverage from patenting the improvement outweighs the original patentee’s sunk-cost leverage. This will usually be the case “when the value of the improvement is large. . . [and] the costs of development is low.” *Id.* at 352-53. Siebrasse and Culver believe that this problem is acute enough to undercut the

blocking patents do not give leverage to consumers or producers. Therefore, by giving leverage to improvers but not to consumers and producers, blocking patents cleverly implement the separation principle.

We have explained that blocking patents gives the improver leverage in negotiating a license fee with the original inventor. If courts wish to increase the improver's leverage further, they can do so in several ways. One way is to rehabilitate the reverse doctrine of equivalences. This obsolete doctrine empowers the court to except radical improvements from liability, even when they literally infringe a previous patent.¹⁵³ By excepting radical improvements, the reverse doctrine of equivalence effectively made radical improvements independent, thereby narrowing the *scope* of the original patent.

Another way for courts to increase the improver's leverage is to weaken the original inventor's *remedy* for infringement. As depicted in Figure 9, damages are a weaker remedy than an injunction.¹⁵⁴ Courts could award damages to the original inventor for infringement by an innovator, not an injunction. Substituting damages for an injunction would effectively create a compulsory license regime.¹⁵⁵

We discussed two ways for patent law to favor the improver over the original inventor: the reverse doctrine of equivalence and substituting damages for an injunction against infringing. When law favors the improver over the original inventor, wealth transfers from the latter to the former. Will this transfer increase the rate of progress

potential value of the experimental use defense (“[W]e conclude that the European [experimental use] approach is not likely to have a major practical effect on the development of improvements. Other considerations, in particular sunk costs incurred in product development, encourage licensing by a potential improver prior to experimentation in any event.”). *Id.* at 334-35. We doubt that this is the case given that experimentation will not necessarily result in a blocking patent. *Cf.* Eisenberg, *supra* note 135, at 1076 (discussing the possibility of a noninfringing substitute).

¹⁵³ See generally Karl Bozicevic, *The “Reverse Doctrine of Equivalents” in the World of Reverse Transcriptase*, 71 J. PAT. & TRADEMARK OFF. SOC’Y 353 (1989); see *supra* note 83.

¹⁵⁴ A threat of an injunction compels the user to pay the owner's asking price, whereas the damage remedy allows the user to choose between buying the right at the owner's asking price or taking it and paying court-set damages. Taking from the owner is involuntary. Thus, allowing the improver to infringe on the original inventor's patent and to pay damages weakens the latter's rights against the former. See *supra* Section IV.C.

¹⁵⁵ See *IP Bargaining Breakdown*, *supra* note 33, at 104 (“Some countries, such as France, Italy, and Japan, grant licenses for all blocking patents; others, such as China and Sweden, reserve the license for cases where the improvement makes a ‘significant technical advance’ over the basic invention, roughly the same as my notion of radical improvements.”).

in the useful arts? The answer depends on their relative fertility. Patent protection that transfers wealth from less to more fertile innovations will increase the pace of innovation according to the fertility principle.¹⁵⁶ So, the court might adopt the rule that *very* fertile improvements do not infringe on the original patent (reverse doctrine of equivalence), or, if they do infringe, damages become the remedy.

The relative fertility of the improvement and the original invention might also impact the optimal patent strength that is granted against the experimentation that led to the improvement. For example, if courts excuse the improvement from liability under the reverse doctrine of equivalents, they might consider rewarding the original patentee damages for the experimentation that led to the improvement so as not to leave the original patentee completely uncompensated.¹⁵⁷

In theory, the court can choose one out of six possible treatments for an experimentation that leads to a fertile improvement:¹⁵⁸ (1) award damages for experimentation + leave the improvement blocked; (2) excuse the experimentation + leave the improvement blocked; (3) award damages for experimentation + award damages for the improvement, (4) excuse the experimentation + award damages for the improvement; (5) award damages for experimentation + excuse the improvement; and (6) excuse both the experimentation and the improvement.

The more fertile the improvement relative to the original invention, the further a court will be willing to go. For example, when a Nobel Prize-winning improvement in the process of making commercial plastic was accused of infringing on an older patent in the field, a court might have found that the improvement's fertility justifies prescribing one of the later treatments in this list.¹⁵⁹

¹⁵⁶ See, e.g., Lemley, *supra* note 25, at 1073 (“[E]fficiency is best served by some sort of calibration, however rough, between the importance of the invention and the scope of the patent.”). Lemley refers to “efficiency.” Our discussion of optimal patent strength implements this proposition with respect to growth, which is sometimes called “dynamic efficiency.”

¹⁵⁷ Cf. Eisenberg, *supra* note 135, at 1078 (suggesting that denying damages in such cases would “lead to unjust enrichment of the researcher”).

¹⁵⁸ This simplified analysis assumes two levers for tailoring patent strength: (1) damages/injunction, and (2) infringement/defense from infringement.

¹⁵⁹ It didn't. This example is based on Phillips Petroleum Co. v. U.S. Steel Corp., 673 F. Supp. 1278 (D. Del. 1987), a case that was part of the Hogan saga, involving the invention of Polypropylene. In this case, the court denied the application of the reverse doctrine of equivalence, which could have excluded the improvement from infringement even though the improvement was presumably very fertile as indicated by the fact that it has resulted in its inventors winning a

Same as noted for experimentation, we do not attempt to strike the perfect degree of patent strength for improvements or suggest that courts should have ultimate discretion tailoring patent strength during litigation.¹⁶⁰ Our focus here is on separation which does not require delicate balancing. From this perspective, it is sufficient to say that the legal framework is only partly in compliance with this paper's thesis. While the doctrine of blocking patents adheres to the separation principle, the reverse doctrine was long made obsolete,¹⁶¹ damages are not usually prescribed for infringement by fertile improvements,¹⁶² and even the patentability of improvements is not usually considered a factor in the infringement analysis.¹⁶³ Instead, as explored next, the legal

Nobel Prize. See Charles Adams, *Allocating Patent Rights Between Earlier and Later Inventions* 54 ST. LOUIS U. L.J. 55, 85 (2009) ("Ziegler and Natta were the first to synthesize polypropylene, and shared the 1963 Nobel Prize for doing so."). The improvement was also extremely successful commercially. In the words of the court, "the extensive array of commercial products paraded into the courtroom by Defendants was most impressive." *Phillips*, 673 F. Supp. at 1357. Nevertheless, commercial success is usually a poor indicator of fertility. Commercial success is measured by the value of the invention to consumers, whereas growth is measured by the value of the invention to innovators and researchers. Thus, the court was correct to discount the evidence of commercial success when evaluating the applicability of the reverse doctrine of equivalences. In the words of the court, "commercial success has no bearing on the Court's infringement determination. Commercial success is simply not the test of infringement." *Id.* See also notes 194–199 and accompanying text.

¹⁶⁰ Greater court discretion to tailor patent strength will undermine the notice function of patents, which might undermine innovation incentives and hinder growth. See Freilich, *supra* note 72 (discussing the tension between actuate incentives and notice); notice function below. Peter S. Menell & Michael J. Meurer, *Notice Failure and Notice Externalities*, 5 J. LEGAL ANALYSIS 1 (2013) (exploring the role of the notice function); *Merrill v. Yeomans*, 94 U.S. 568, 573 (1876) ("[P]atent law . . . leave[s] no excuse for ambiguous language or vague description."); *McClain v. Ortmayer*, 141 U.S. 419, 424 (1891); *Halliburton Energy Servs., Inc. v. M-I LLC*, 514 F.3d 1244, 1249 (Fed. Cir. 2008).

¹⁶¹ See *supra* notes 83 and 153.

¹⁶² Indeed, most commenters interpreted *eBay* to address concerns about holdup and uncertain rights. See *infra* notes 192, 170 and accompanying text. See also BURK & LEMLEY, *supra* note 9, at 140 ("[I]njunctive relief may be inappropriate where patent rights are asserted primarily as holdups rather than as part of an effort to protect a legitimate invention."); Menell & Meurer, *supra* note 160, at 52 ("The awarding of damages presents a useful place to integrate concerns about notice externalities. There is good to reason to discount damage awards in those circumstances in which it is particularly difficult to identify and evaluate intellectual property encumbrances."); Carl Shapiro, *Property Rules vs. Liability Rules for Patent Infringement* (manuscript at 25), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=27-75307 ("If these switching costs are small relative to the value of the patented technology, an injunction is likely to be the better remedy.").

¹⁶³ But see *Zygo Corp. v. Wyko Corp.*, 79 F.3d 1563 (Fed. Cir. 1996) (suggesting that an improvement's patentability is evidence of noninfringement); Merges &

discourse on weakening patents' strength has recently moved away from the separation function to other ancillary objectives. This trend has misdirected patent policy away from the constitutional mandate.

B. Weakening Patents Against Production

As discussed, weakening patent strength against innovation generally speeds up growth. In special circumstances, however, the opposite is true: Weakening patent strength against production speeds up growth. These circumstances apply when bargaining over licensing fails. Bargaining failure has two main causes: holdout and uncertain rights.

However, weakening patent strength to ameliorate bargaining failures is not a fundamental function of patent law. Overstating such policies undermines separation and may deteriorate growth. Thus, policy makers should use caution when they weaken patent strength against production. Policy makers should then favor policies that mitigate bargaining failure with a minimal impact on patent strength.

Alas, in recent years, inflated concerns about holdout and uncertain patent rights have taken center stage, as well as confused and overshadowed the separation function of patents. Below, we clarify these issues and properly situate them as secondary considerations.¹⁶⁴ In Section I, we explore the holdout problem, and in Section ii, we explore the problem of uncertain rights.

i. Holdout

Holdout is a problem of collective action that prevents economic players from reaching an efficient bargain outcome. The classic example assumes that the state wants to build a road through five parcels of private land: $A \rightarrow B \rightarrow C \rightarrow D \rightarrow E$. As the state assembles land parcels, it commits to this route. Buying four parcels, say (A,B,C,E), locks the state into buying parcel D. By refusing to sell, the owner of parcel D can stop the entire project. Every property owner recognizes that a parcel becomes more valuable to the state after it purchases the others along the route. Given this logic, every

Nelson, *supra* note 33, at 909 (arguing that courts should consider the nature of the allegedly infringing use as part of the infringement analysis).

¹⁶⁴ By classifying the principles of separation and fertility as “primary” and the problems of holdout and uncertain rights as “secondary,” we do not mean to imply that the former is necessarily more important than the latter. Instead we wish to emphasize that while the principles of separation and fertility are fundamental and endemic to patent policy, the problems of holdout and uncertain rights are generic problems of property law. As such these problems can often be addressed by subtler or non-patent measures. See *infra* notes 171–176, 201–207 and accompanying text.

property owner wants to sell last. If every seller holds out, the road will never get built.

The same holdout problem arises when producing goods from several patented technologies.¹⁶⁵ Assume, for example, that manufacturing a new device—say, a mobile phone—requires five patented inputs $(\alpha, \beta, \chi, \delta, \varepsilon)$.¹⁶⁶ Each one is essential to the device, and each one has a different owner. Buying four licenses, say $(\alpha, \beta, \chi, \varepsilon)$, locks the manufacturer into buying a license for δ to complete the device. Since withholding license δ stops the entire project, its owner can extract a high price. Every patent owner recognizes this logic, so every patent owner wants to sell last. If every seller holds out, the device will never get built.

The risk of holdout is very real in the patent arena, especially in industries where research and development is a cumulative process that requires accumulation of multiple patented inputs from various right holders.¹⁶⁷ Nevertheless, in some instances, competition solves the holdout problem. When the owner of an input holds out for a higher price, a competitor eventually enters the market and makes a substitute. With patent holdout, substitution requires inventing around the patent or waiting until it expires. A single substitute weakens the power to hold out, and many substitutes eliminate it. Thus, competition eventually solves some holdout problems in private business.

However, this is not always true. In some cases, patented inputs are essential, not substitutable. In such cases, competition cannot resolve the holdout problem. As with the state building a road, patented inputs become essential by the choice of the path of research or development. To illustrate in our notation, acquiring $(\alpha, \beta, \chi, \varepsilon)$ locks the manufacturer into acquiring δ . Relative to development, δ was minor ex-ante and essential ex-post.

Why would a manufacturer choose a development path with essential patents owned by others? Perhaps he did not know that δ

¹⁶⁵ See generally MANCUR OLSON, *THE LOGIC OF COLLECTIVE ACTION* (Harvard Univ. Press, 1961). On its application in patent law, see Rochelle Cooper Dreyfuss, *Varying the Course in Patenting Genetic Material: A Counter-Proposal to Richard Epstein's Steady Course*, in *PERSPECTIVES ON PROPERTIES OF THE HUMAN GENOME PROJECT* 196 (F. Scott Kieff ed., 2005).

¹⁶⁶ Mark A. Lemley & Carl Shapiro, *Patent Holdup and Royalty Stacking*, 85 TEX. L. REV. 1991 (2007) (“As a striking example, literally thousands of patents have been identified as essential to the proposed new standards for 3G cellular telephone systems.”).

¹⁶⁷ Carl Shapiro, *Navigating the Patent Thicket: Cross Licensing, Patent Pools, and Standard Setting*, in *INNOVATION POLICY AND THE ECONOMY* 119 (Adam Jaffe et al. eds., 2001) (focusing his analysis on cumulative innovation industries “including semiconductors, biotechnology, computer software, and the Internet”).

would be necessary until development commenced. Or, perhaps he knew that δ would be necessary, but he did not know that it was patented. As explained in the next section, in the realm of patents, information asymmetries are acute.¹⁶⁸

When holdout is substantial, efficient transactions are defeated, patent holders are prevented from maximizing their profits from innovation, and economic growth deteriorates. In such cases, policy makers can weaken patent rights against production and overcome the holdout problem. Weakening patents' remedy from an injunction to damages is the most cited example. As explained in Part IV.C, a damage remedy allows users to take the patented input (and pay the court-set price) rather than buying the input and paying the owner's asking price. As with the state failing to build a highway, holdout in private bargaining requires *taking* the patented input, not *buying* it.¹⁶⁹ Thus, some interpret the recent Supreme Court decision in the case of *eBay* to be a useful policy lever to ameliorate holdout.¹⁷⁰

We do not contest the logic of this approach, but we do contend that it should be limited in application and not overstated. Weakening patents against production undermines the separation principle, and if used in excess, it would reduce profitability from innovation and decrease economic growth. In recent years, inflated

¹⁶⁸ See Mark A. Lemley, *Ignoring Patents*, 19 MICH. ST. L. REV. 20, 28 (2008) ("It's not that nothing of that sort ever happens with land; commercial builders often face a holdout problem when they need to buy numerous plots of contiguous land for a new building. But the problem is much, much worse in the patent system because the number of rights that must be aggregated is greater, the scope of those rights (and therefore which ones need to be included) is less clear, and courts may be unable to grant injunctive relief tailored to protect only the patent in question without interfering with the non-infringing uses. The result is that bargaining breakdown, already a risk in real property (and the justification for the use of eminent domain in support of private projects), will be much more of a risk."); Menell & Meurer, *supra* note 160 (explaining the uncertainty is much worse in intangible property relative to tangible property).

¹⁶⁹ Cf. U.S. CONST. amend. V. ("[N]or shall private property be taken for public use, without just compensation.").

¹⁷⁰ See *eBay Inc. v. MercExchange, L.L.C.*, 547 U.S. 388, 396 (2006) (Kennedy, J., concurring) (stating that "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" and proposing that if "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"); Gavin D. George, *What Is Hiding in the Bushes?: eBay's Effect on Holdout Behavior in Patent Thickets*, 13 MICH. TELECOMM. & TECH. L. REV. 557, 566 (2007) ("Most importantly, this decision has the potential to provide an effective defense against patent holdouts seeking to recover switching costs.").

concerns about holdout led to aggressive policies that weakened patent protection beyond the spirit of the separation mandate.¹⁷¹

For as long as other means are available to mitigate bargaining failure with minimal impact on patent strength, such policies should be favored. Prominent scholars, such as Carl Shapiro and Robert Merges, have systematically argued that producers often consolidate and contract their way out of the holdout predicament voluntarily.¹⁷²

For example, the contract that creates most standard-setting organizations requires its members to stipulate a damage remedy for infringement, not an injunctive remedy. Such a contract creates a regime of private takings with compensation. A common stipulation of damages is “fair, reasonable, and non-discriminatory” (FRAND or RAND), which lets the court set the specific value of damages.¹⁷³ As Professor Merges argued, such arrangements are facilitated precisely by strengthening patent entitlements, not by weakening them.¹⁷⁴

However, not all holdout scenarios contradict the separation principle. For example, a medical scientist may need several patented devices to experiment with diagnostic tools. If each patented device becomes more valuable to the experimenter after licensing the others, each owner of the patented devices may hold

¹⁷¹ Some critics have stated that the vague guidance of eBay has undermined the faith in the patent system and led to “efficient infringement,” where users opt to infringe and pay compensation later only if they get caught. *See, e.g.,* Chris Gallagher & Gene Quinn, *Move Over Patent Trolls, Efficient Infringement Has Arrived on the Hill*, IPWATCHDOG (Oct. 25, 2016), <https://www.ipwatchdog.com/2016/10/25/efficient-infringement-arrived-hill/id=74131/>.

Others feel that the overstated fears about holdout and uncertain rights gave birth to the *Alice-Mayo* framework to patent eligibility, which is heavily criticized. *See supra* notes 65–68 and the accompanying text.

¹⁷² *See* Shapiro, *supra* note 167; Robert P. Merges, *Contracting into Liability Rules: Intellectual Property Rights and Collective Rights Organizations*, 84 CALIF. L. REV. 1293 (1996).

¹⁷³ These terms are often vague, which causes disputes about how to compute reasonable damages. Adoption of a patent as a standard greatly increases its value. Therefore, an interesting question is, should damages for infringing on the patent include or exclude the value added to it by its adoption as a standard? For analysis, *see* Richard Gilbert, *Deal or No Deal? Licensing Negotiations by Standard Development Organizations*, 77 ANTITRUST L.J. 855, 855–88 (2011). Also, contracting for standards often requires negotiating among competitors in an industry, which might create antitrust concerns.

¹⁷⁴ Merges, *supra* note 172, at 1303 (“[T]he property rule entitlements granted at the outset actually lead to a liability rule-like regime, though one based on collective valuation by firms, rather than by an arm of government.”).

out for a high price.¹⁷⁵ Thus, when holdout blocks innovation (as opposed to production), weakening patent rights to curtail holdout falls squarely within the separation principle.¹⁷⁶

Gene patents, for example, are often claimed to burden biomedical research. To deal with this problem, various scholars have suggested to narrow the *breadth* of such patents¹⁷⁷ or to weaken the *remedy* that would be prescribed when these patents are infringed.¹⁷⁸ Even though these commenters often frame these suggestions as a means to ameliorate holdout, they essentially promote separation by weakening patent strength against innovation.

ii. *Uncertainty*

We began this section by explaining that strong patent rights sometimes cause bargaining failures that reduce the overall profitability of innovating. We have been discussing the first cause of bargaining failure, “holdout,” which involves several different sellers of essential inputs. Now we turn to the second cause of bargaining failure, which is uncertainty about property rights. Uncertain rights can cause bargaining failure even with only one seller and one buyer. The problem of uncertainty is logically distinct from the holdout problem.

We already noted that information asymmetries are acute in patents. With undecipherable records and outdated secrecy rules, it is difficult to identify all issued or pending patents.¹⁷⁹ The validity of issued patents is doubtful until they are litigated;¹⁸⁰ many will

¹⁷⁵ Cf. Dreyfuss & Evans, *supra* note 71, at 1365 (“[G]enetic research requires broad access to both phenotype and genotype information.”); Michael A. Heller, *The Tragedy of the Anticommons: Property in the Transition from Marx to Markets*, 111 HARV. L. REV. 621 (1998); Michael A. Heller & Rebecca Eisenberg, *Can Patents Deter Innovation? The Anticommons in Biomedical Research*, 280 SCI. 698, 698–99 (1998).

¹⁷⁶ Nevertheless, to avoid confusion, it helps to keep the issues distinct. Patent strength should be weakened against innovation even in the absence of holdout.

¹⁷⁷ See, e.g., Heller & Eisenberg, *supra* note 175; See also BURK & LEMLEY, *supra* note 9, at 77 (“Most legal scholars working in the anticommons literature have assumed that the solution is to grant fewer patents, particularly to developers of upstream products such as research tools or DNA sequences.”); Assoc. for Molecular Pathology v. Myriad Genetics, Inc., 569 U.S. 576 (2013).

¹⁷⁸ See, e.g., Donna M. Gitter, *International Conflicts over Patenting Human DNA Sequences in the United States and the European Union: An Argument for Compulsory Licensing and a Fair-Use Exemption*, 76 N.Y.U. L. REV. 1623 (2001) (suggesting a compulsory license regime for DNA sequences to solve holdout).

¹⁷⁹ See generally Menell & Meurer, *supra* note 160.

¹⁸⁰ Mark A. Lemley & Carl Shapiro, *Probabilistic Patents*, 19 J. ECON. PERSP. 75 (2005) (“A patent does not confer upon its owner the right to exclude but rather a right to try to exclude.”); Mark A. Lemley, *Rational Ignorance at the Patent*

never be litigated, and of those litigated, almost 50% will be invalidated.¹⁸¹ Unlike tangible property such as land, the legally protected boundaries of patents are inherently vague, especially in information technology.¹⁸² The scope of individual patent rights necessarily overlaps to create a “thicket.” Fuzzy boundaries, infeasible clearance, and treble damages for willful infringement incentivize developers to ignore patents altogether.¹⁸³

When rights are unclear, developers might accidentally encroach upon patented territory without knowing that relevant patent rights exist. In these circumstances, developers would have to bargain for a license after the investment in production was made rather than before. At this stage, even if the patented input is not essential for production, strong patent rights give patent owners a substantial strategic advantage by threatening infringement. With an injunction threat looming to shut down its production line, a rational producer might prefer overpaying for a license rather than redesigning production to avoid infringement or litigating to prove noninfringement.¹⁸⁴

Even when developers can recognize patents before investing in production, uncertain rights may still frustrate bargaining. Private bargaining fails because unclear rights create scope for false optimism about litigation. With false optimism, both sides expect to win in litigation, so they fail to reach an

Office, 95 NW. U. L. REV. 1, 2 (2001) (explaining that “the PTO doesn’t do a very detailed job examining patents”); Hacoheh, *supra* note 14 (exploring why patents are inevitably probabilistic).

¹⁸¹ John R. Allison & Mark A. Lemley, *Empirical Evidence on the Validity of Litigated Patents*, 26 AIPLA Q.J. 185 (1998).

¹⁸² See Menell & Meurer, *supra* note 160; BURK & LEMLEY, *supra* note 9, at 138 (“Closely related to the problem of complementarity is the problem of horizontal overlaps between patents.”); Shapiro, *supra* note 167.

¹⁸³ Reading patents may signal to the courts that the infringement was willful, which will result in treble damages. 35 U.S.C. § 284 (2018). See generally Lemley, *supra* note 168. In *Halo Elec., Inc. v. Pulse Elec., Inc.*, 136 S. Ct. 1923 (2016), the Supreme Court increased the likelihood that the courts could award triple damages that could potentially enhance nuisance litigation. See Chase Means, *Has the Supreme Court Breathed New Life into Patent Trolls in Halo and Stryker?*, IPWATCHDOG (June 15, 2016), <https://www.ipwatchdog.com/2016/06/15/supreme-court-patent-trolls-halo-stryker/id=70050>. Moreover, even the most thorough search will not be able to reveal pending patent application at the patent office. Strategic use of continuation practice to keep patents pending and hidden longer at the PTO has become an increasingly common and criticized practice. FED. TRADE COMMISSION, THE EVOLVING IP MARKETPLACE: ALIGNING PATENT NOTICE AND REMEDIES WITH COMPETITION 80–90 (2011); Richard A. Posner, *Why There Are Too Many Patents in America*, ATLANTIC (July 12, 2012), <http://www.theatlantic.com/business/archive/2012/07/why-there-are-too-many-patents-inamerica/259725>; Hacoheh, *supra* note 14.

¹⁸⁴ For analysis, see Shapiro, *supra* note 167; Menell & Meurer, *supra* note 160.

agreement.¹⁸⁵ In such circumstances, courts may need to clarify the rights before private bargaining can succeed. However, clarifying uncertain rights through legal processes takes time. A manufacturer who needs to buy a license often suffers more from delay in clarifying the rights than its owner.¹⁸⁶ So, the plaintiff can threaten a suit, and a rational defendant will pay to settle it.

Lawsuits that are brought to induce a settlement, not to win at trial, are called nuisance (or negative value) lawsuits. In such cases, the plaintiff does not sue to win at trial because his trial costs exceed his expected benefit from the court's judgment. The classic example of a nuisance suit involves building permits in New York City. Modern building in cities requires many permits issued in stages as construction proceeds. Neighbors and other interested parties can often delay construction by suing for an injunction. So, the builder often pays to settle a case that he would eventually win at trial.¹⁸⁷ In patent cases, the costs and uncertainty of patent litigation as well as the asymmetrical costs of delay are basic causes of nuisance lawsuits.

Nuisance lawsuits in the patent arena are often pursued by patent assertion entities (PAEs) or pejoratively "patent trolls." These entities specialize in licensing patents by threatening to sue infringers, not inventing or producing. PAEs were responsible for roughly two-thirds of all patent litigation cases brought in the U.S. in 2015.¹⁸⁸ Since litigation is their specialty, PAEs have advantageous bargaining positions relative to defendant manufacturers.¹⁸⁹ They do not fear production delays because they produce nothing.¹⁹⁰ They are accused of aggressive litigation and

¹⁸⁵ For a summary of litigation bargaining, with illustrations of false optimism, see COOTER & ULEN, *supra* note 97, at ch. 10.

¹⁸⁶ In technical terms, asymmetrical litigation costs imply that the Bayesian bargaining solution is positive, even though the expected judgment is less than the plaintiff's cost of litigating. Note that nuisance suits mostly occur when each side pays his own litigation costs, not when the loser pays the winner's litigation costs. *See Id.*

¹⁸⁷ To solve the hold-up problem in construction, private builders often partner with state authorities, who can use their legal powers to overcome permitting problems. By paying for permits, the builders may be able to avoid paying the price demanded by neighbors who hold them up.

¹⁸⁸ *See* Mackenzie Weinger, *Patent Trolls Rear Their Ugly Heads in Courtrooms Around the World*, FIN. TIMES (June 9, 2016), <https://www.ft.com/content/a1427d4c-0e25-11e6-b41f-0beb7e589515>.

¹⁸⁹ *See* Brian T. Yeh, *An Overview of the "Patent Trolls" Debate*, DIG. LIBRARY 12 (Apr. 16, 2013), <https://digital.library.unt.edu/ark:/67531/meta-dc462666>.

¹⁹⁰ *Id.* at 13 ("Because PAEs . . . do not make or sell anything, they are not subject to counterclaims that they infringe on defendants' patents. . . . Compounding that leverage, the PAE business model creates unusual incentives for PAEs to forge ahead with weak suits (rather than calling it a loss or accepting a lowball

nuisance suits. Some research suggests that “[venture capital] investment would have likely been \$21.772 billion higher over the course of five years but-for litigation brought by frequent litigators.”¹⁹¹

As in holdout, weakening patents’ strength against any use—consuming, producing, or innovating—would overcome the problems caused by uncertain rights. Uncertain rights breed legal disputes and weakening rights dampen legal disputes. Thus, many have argued that, following the *eBay* decision, courts should reduce the *remedy* for infringement from injunction to damages to curtail nuisance lawsuits advanced by NPEs.¹⁹² Other scholars argued that patent *breadth* should also be narrowed to curtail nuisance lawsuits. Professor Ernst, for example, suggested that the reverse doctrine of equivalence should be interpreted to exclude uses in production when patents are enforced by NPEs.¹⁹³

These developments are concerning because they deflect the conversation away from the patent system’s fundamental function—separation—while overemphasizing ancillary considerations such as successful commercialization of inventions and overvilifying NPEs. Commenters have long warned that commercialization is a poor measure for innovative advancement.¹⁹⁴ Still, the role of commercial success in patent doctrines has only heightened over

settlement) to reinforce their bargaining position with future targets. The ability to extract licensing fees depends upon posing a credible threat of costly litigation.”).

¹⁹¹ Stephen Kiebzak et al., *The Effect of Patent Litigation and Patent Assertion Entities on Entrepreneurial Activity*, 45 RES. POL’Y 218, 229 (2016).

¹⁹² *eBay Inc. v. MercExchange, L.L.C.*, 547 U.S. 388, 395-96 (2006) (Kennedy, J., concurring) (“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.”); BURK & LEMLEY, *supra* note 100, at 140 (“[I]njunctive relief may be inappropriate where patent rights are asserted primarily as holdups rather than as part of an effort to protect a legitimate invention.”); Shapiro, *supra* note 162, at 25 (“Our analysis indicates that the court should carefully assess the switching costs that the infringing party would incur if forced to stop selling infringing products. . . . If these switching costs are small relative to the value of the patented technology, an injunction is likely to be the better remedy.”).

¹⁹³ Samuel F. Ernst, *The Lost Precedent of the Reverse Doctrine of Equivalents*, 18 VAND. J. ENT. & TECH. 367, 475–76 (2016) (recommending that “in addition to examining the accused product’s technological change from the patented invention, courts should also compare the commercial success of the accused product to any product manufactured by the patent-holder.”). Professor Ernst rightly acknowledged that commercial success may also skew the analysis and thus aptly demand that the accused infringer would need to show “a nexus between the required commercial success of the accused product and the substantial technological improvement it offers.” *Id.*

¹⁹⁴ See, e.g., Robert P. Merges, Commercial Success and Patent Standards: Economic Perspectives on Innovation, 76 CALIF. L. REV. 805, 838-39 (1988).

time. Most troublingly, in its 2016 en banc decision in the case of *Apple v. Samsung*,¹⁹⁵ the Federal Circuit had “turn[ed] obviousness analysis on its head”¹⁹⁶ by grossly overstating the role of commercial success as an indicia for inventiveness.¹⁹⁷ In one noticeable example, the majority relied on the audience’s cheers in an Apple event as an indication that Apple’s slide-to-unlock feature was nonobvious.¹⁹⁸ However, as explained in Part III.B, inventiveness is measured in the market of innovators, not the market of consumers.¹⁹⁹

Similarly, just as commercialization should not be praised in patent policy, a lack of commercialization should not be disparaged. By connecting buyers and sellers, NPEs generally perform a valuable social service of moving goods, including patents, to their highest valued use. By doing so, they benefit the makers and consumers of goods, including patents.²⁰⁰

We do not mean to suggest the “trolling” concerns are unimportant or that they should not be addressed in patent policy. To the contrary, a host of commenters have rightly argued that trolling behavior imposes an unnecessary drag on innovation and thus impedes economic growth.²⁰¹ We do suggest, however, that by carefully tracing the roots of the trolling phenomenon policymakers

¹⁹⁵ 839 F.3d 1034 (Fed. Cir. 2016) (en banc).

¹⁹⁶ Petition for Writ of Certiorari, *Samsung Elec. Co. vs. Apple Inc.*, 138 S. Ct. 420 (2017) (No. 16-1102) (cert. denied).

¹⁹⁷ *Apple Inc. v. Samsung Elec. Co., Ltd.*, 839 F.3d 1034, 1080 (Fed. Cir. 2016) (Dyk, J., dissenting) (“[T]he majority errs in elevating secondary considerations of nonobviousness beyond their role as articulated by the Supreme Court.”).

¹⁹⁸ *Id.* at 1054 (Moore, J.).

¹⁹⁹ *Id.* at 1083 (Dyk, J., dissenting) (noting that Apple provided no evidence that such praise was directed specifically for the slide-to-unlock feature or that the audience was comprised of industry experts).

²⁰⁰ In general, patent intermediaries enforce the patents that they buy, which increases their effective strength. Increasing the effective strength of patents can increase or decrease the rate of innovation. The separation principle formulates the most basic difference. Increasing the effective strength of patents against consuming and producing causes faster growth and increasing the effective strength of patents against innovating usually causes slower growth. Therefore, patent intermediaries cause faster growth by enforcing patents against consuming and producing, and intermediaries usually cause slower growth by enforcing patents against innovating. We thus suggest reserving the pejorative term “patent trolls” for intermediaries who decrease innovation. Intermediaries usually decrease innovation by suing innovators for infringing or by suing producers; doing so aggravates a holdout problem or vague rights.

²⁰¹ See Shapiro, *supra* note 167 at 121 (labeling holdup as a tax on innovation); James Bessen & Michael J. Meurer, *The Direct Costs from NPE Disputes*, 99 CORNELL L. REV. 387, 417 (2014) (same); Menell & Meurer, *supra* note 160, at 39.

can come up with more nuanced policy measures that address the trolling problem with a minimal impact on patent strength.²⁰²

Professors Peter Menell and Michael Meurer, for example, have explored various institutional and procedural inefficiencies at the USPTO that unnecessarily aggravate patent uncertainty.²⁰³ In accordance, they advance a variety of remedial proposals—progressive funding mechanism, improved quality review procedures, transparency in the application filings, and broader public participation in the examination process—that would all greatly improve patent certainty without undermining patent strength.²⁰⁴ In a similar vein, because nuisance lawsuits thrive on the high costs of patent litigation, post-grant remedial measures—improved opposition procedures,²⁰⁵ heightened litigation sanctions,²⁰⁶ and fee-shifting²⁰⁷—would also ameliorate the trolling problem with a minimal impact of patent strength.

²⁰² Some degree of patent uncertainty is nevertheless inevitable. Indeed, patent policy cannot coherently conform to the principles of separation and fertility without contributing to the problem of uncertain rights. For example, by requiring to weaken patent strength against innovative uses, the separation principle aggravates the uncertainty surrounding patent scope and encourages nuisance lawsuits. *See generally* Freilich, *supra* note 72 (describing the inevitable tension between accurate patent incentives and clear patent boundaries). In the limited scope of this Article we do not intend to craft the optimal balance between accuracy and certainty, but we do emphasize that patent doctrines are insufficiently sensitive to the former and disproportionately attentive to the latter.

²⁰³ *See generally* Menell & Meurer, *supra* note 160 (describing in great detail the reasons for patent uncertainty in the patent arena).

²⁰⁴ *Id.* at 29.

²⁰⁵ *See* Yeh, *supra* note 189, at 12 (“The AIA provisions increasing the speed and availability of post-grant examination is expected to ameliorate this issue somewhat for invalid patents granted after 2011. Defendants will be able to challenge a patent’s validity, but not its scope or the claim of infringement, at a much lower cost than they can in court, where they must overcome a presumption of validity by clear and convincing evidence to get a patent invalidated.”). *But see* Michael Gulliford, *If Patent Reform Is Meant to Starve Patent Trolls, Why Is It Feeding Them Instead?*, IPWATCHDOG (Sept. 8., 2014), <http://www.ipwatchdog.com/2014/09/08/if-patent-reform-is-meant-to-starve-patent-trolls-why-is-it-feeding-them-instead/id=51067>.

²⁰⁶ *See, e.g.,* Gene Quinn, *Indicia of Extortion—Federal Circuit Slams Patent Troll*, IPWATCHDOG (Aug. 4, 2011), <https://www.ipwatchdog.com/2011/08/04/indicia-of-extortion-federal-circuit-slams-patent-troll/id=18568>.

²⁰⁷ *See* Christian Helmers, Brian Love & Luke McDonagh, *Is There a Patent Troll Problem in the UK?*, 24 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 509, 516 (2015) (“Our findings tend to suggest, instead, that attorney’s fee awards are a key factor in the scarcity of NPE activities in Europe.”); Rudy Telscher & Kara R. Fussner, *Patent Fee Shifting Stops Not Only Patent Trolls But Industry Bullies Too*, IPWATCHDOG (Sept. 16, 2015), <http://www.ipwatchdog.com/2015/09/16/patent-fee-shifting-stops-not-only-patent-trolls-but-industry-bullies-too/id=61575>.

VI. CONCLUSION

The U.S. Constitution authorizes Congress to issue patents to promote progress in the useful arts, which we interpret as increasing economic growth through innovation. Innovation refers to discovering better goods to make or better techniques to make them. To ground patent law, we formulate two principles of growth economics. First, patent law increases profits for innovation by increasing the price of consuming and producing the innovation. Transferring wealth from consuming and producing to innovating accelerates innovation. To speed up growth, patent protection should be strong against using an innovation to produce or consume.

Conversely, when innovators buy and sell patents among one another, they gain less overall as sellers than they lose as buyers (“deadweight loss”). Redistributing wealth among innovators usually slows innovation. To avoid slowing innovation, patent protection should be weak against using an innovation to innovate. Combining the two preceding propositions implies strengthening patents against consuming and producing and weakening patents against innovating (separation principle).

Second, human welfare can increase exponentially from innovation. With exponential growth, the increase in human welfare quickly overtakes any losses from static inefficiency or inequality (overtaking principle). Welfare overtaking is the ethical and political justification of the Constitution’s patent clause.

A patent restricts competition to produce the innovation, which raises its price. A higher price for consumers and producers causes more innovation. However, a higher price than the cost of production is inefficient. Also, the price affects the distribution of income. The standard evaluation of patent law balances the benefits of growth, the cost of static inefficiency, and the cost of any increase in inequality. Balancing is scientifically difficult and philosophically controversial.

With rapid innovation, growth in human welfare overtakes losses from inefficiency or inequality. The effects of patents on efficiency and distribution on welfare are unimportant relative to their effects on growth. Like the constitution, welfare overtaking suggests that patent interpretation and policy should focus on innovation and dispense with balancing.

Separation and welfare guide patent law toward its constitutional purpose. To achieve growth, the law implements separation, by weakening patents against innovating without weakening patents against consuming and producing. We have illustrated that these principles should guide policy through legal doctrines such as the experimental use defense, blocking patents,

and private taking of patents. We also clarify exceptions to these principles in circumstances of bargaining failures but properly situate these considerations as secondary.