

## SPILOVERS

Brett M. Frischmann\* & Mark A. Lemley\*\*

*This Essay challenges the conventional law and economics account of intellectual property. The Essay combines supply-side and demand-side considerations and develops a theory that leads to important insights about the relationship between spillovers (positive externalities) and law generally. Contrary to the commonplace “law and economics” view of spillovers as a problem in need of solution, it turns out that a wealth of economic evidence teaches us that spillovers are good for society, and in many different contexts the law is actually designed to encourage certain externality-producing activities. This new theory has significant potential as an alternative economic theory to understand IP as well as other areas of the law. The Essay explores a few of the many possible applications of the theory to copyright and patent law and to the debate over network neutrality. It concludes with some general thoughts about the Coasean symmetry between externalities and property rights. When relevant, both externalities and property rights distort the market allocation of resources; when irrelevant, neither does. Courts and scholars must resist the easy answer of equating public and private value by internalizing externalities because spillovers aren’t always bad, and more property rights aren’t always good. Only if we understand when and why each can enhance social welfare can we hope to design legal rules that do more good than harm.*

### INTRODUCTION

Economists since Demsetz have viewed property rights as a way to internalize the external costs and benefits one party’s action confers on another. They have thought this internalization desirable, reasoning that if a party didn’t capture the full social value of her actions she wouldn’t have optimal incentives to engage in those actions. Measured by this standard, intellectual property (IP) rights are inefficiently weak. There is abundant evidence that the social value of innovations far exceeds the private value. But there is also good evidence that, contrary to what econ-

---

\* Assistant Professor of Law, Loyola University Chicago.

\*\* William H. Neukom Professor of Law, Stanford Law School; of counsel, Kecker & Van Nest LLP.

Thanks to Michael Abramowicz, Mike Carroll, Julie Cohen, Richard Epstein, Joe Farrell, Paul Geller, Shubha Ghosh, Wendy Gordon, Rose Hagan, Paul Heald, Bob Hunt, Larry Lessig, Gary Libecap, Mike Madison, David McGowan, Adam Mossoff, Giovanni Ramello, Matthew Sag, Mark Schultz, Carl Shapiro, Chris Sprigman, Spencer Waller, Phil Weiser, and the participants in the American Law and Economics Association annual meeting, the Georgetown Intellectual Property Colloquium, the Works-in-Progress Intellectual Property Colloquium 2005, cohosted by Washington University and Saint Louis University, and the Working Group on Property, Citizenship, and Social Entrepreneurism, Washington, D.C., November 4–5, 2005, for comments on this or related papers.

omists might assume, these spillovers actually encourage greater innovation. The result is a puzzle for property rights theorists.

In this Essay, we offer three insights that help to explain the positive role of innovation spillovers. First, we explain that in IP, unlike real property, a wide range of externalities matter, because IP rights are much less certain than property rights and because the decision to create a legal entitlement will determine whether or not a transaction must occur. Second, we suggest that there is no reason to think that complete internalization of externalities is necessary to optimize investment incentives; at some point, there are decreasing returns (in terms of improved incentives) to allowing property owners to capture more of the value from their inventions. Spillovers do not always interfere with incentives to invest; in some cases, spillovers actually drive further innovation. Third, we observe that even where internalizing externalities increases incentives to invest, the social costs of relying on property rights to do so still may exceed the benefits. In particular, we focus on demand-side market failures that occur when a potential buyer wants to make productive (re)use of a work. Productive users' demand (willingness to pay) for access to and use of works often understates societal demand because productive users do not capture—and therefore do not take into account—the full social value of their use. Under such circumstances, property rights may skew incentives and create problematic distortions in resource allocation. The result of combining these insights is that, at least where innovation is concerned, we cannot rely on the easy equation of property rights with efficient internalization of externalities.

## I. THE PUZZLE OF SPILLOVERS

### A. *The Nature and Ubiquity of Spillovers*

Spillovers—uncompensated benefits that one person's activity provides to another—are everywhere.<sup>1</sup> My decision to plant an attractive flower bed benefits my neighbors and people who happen to stroll by. Your decision to educate your children well, making them into productive, taxpaying, law-abiding members of society, benefits the people who buy the goods they will produce, the people who will receive the government benefits their taxes fund, and the people they might otherwise have robbed. Dan Bricklin's invention of a computer spreadsheet benefited not just his customers, who paid for the privilege, but their customers, who used spreadsheets; Bricklin's competitors, who built on his idea; and all the people who got products somewhat cheaper because the business from which they bought ran more efficiently by using a computer spreadsheet.

---

1. Cf. Richard A. Epstein, *The Allocation of the Commons: Parking on Public Roads*, 31 *J. Legal Stud.* S515, S520 (2002) (“[I]t is widely understood today that the creation of positive externalities is common under a system of private property.”).

Spillover benefits aren't intentionally provided. We don't plant flowers to benefit our neighbors, you don't educate your children in order to give the government more money, and Dan Bricklin certainly didn't develop the spreadsheet in order to benefit the competitors his invention would eventually attract. These benefits are not the result of altruism. Instead, they are incidental "extras"—they spill over to others as a result of decisions you and we and Bricklin made for our own purposes.<sup>2</sup> While seemingly insignificant to us, spillovers turn out to be enormously significant to society. We are all incidental beneficiaries, each and every day. Spillovers "are a ubiquitous boon for society" because we share a common environment, live in communities, and interact with one another.<sup>3</sup>

The social importance of spillovers is recognized in many contexts, but may be most significant in the context of innovation.<sup>4</sup> Historically, no innovator has captured all or even most of the social benefit of her invention.<sup>5</sup> Her invention might make her some money, but if it does, it will likely make money for other people as well. These spillovers come in

2. See Richard Cornes & Todd Sandler, *The Theory of Externalities, Public Goods, and Club Goods* 55 (1996) ("[T]he only motive that an individual has to provide units of such a good is his or her own private motive of present or future consumption. Enjoyment of those units by others is an incidental by-product.").

3. Brett M. Frischmann, *An Economic Theory of Infrastructure and Commons Management*, 89 *Minn. L. Rev.* 917, 967 (2005) [hereinafter Frischmann, *Economic Theory*] ("Neither the law nor economic efficiency require complete internalization; external benefits are a ubiquitous boon for society.").

4. We use the term innovation here in a way that is broader than invention alone. We mean the term to refer to the process of research, invention, and development and refinement of new ideas.

5. See, e.g., William J. Baumol, *The Free Market Innovation Machine* 5 n.4 (2002) ("[S]urely no major invention has provided benefits only to its inventor."). Put another way, the social returns to private research and development exceed the private return by a substantial margin—a spillover gap. See, e.g., W. Keller & S. Yeaple, *Multinational Enterprises, International Trade, and Productivity Growth: Firm-Level Evidence from the United States* (Leverhulme Ctr., Research Paper No. 03, 2003), available at [http://www.nottingham.ac.uk/economics/leverhulme/research\\_papers/03\\_03.pdf](http://www.nottingham.ac.uk/economics/leverhulme/research_papers/03_03.pdf) (on file with the *Columbia Law Review*) (documenting significant intra-industry spillovers in United States). The following table highlights the results of a few, of the many, econometric studies that have found significant spillovers from private research and development.

PRIVATE AND SOCIAL RATES OF RETURN TO PRIVATE R&D

Author (year)	Estimated Rates of Return	
	Private	Social
Nadiri (1993)	20–30	50
Mansfield (1977)	25	56
Terleckyj (1974)	29	48–78
Sveikauskas (1981)	7–25	50
Goto-Suzuki (1989)	26	80
Bernstein-Nadiri (1988)	10–27	11–111
Scherer (1982, 1984)	29–43	64–147
Bernstein-Nadiri (1991)	15–28	20–110

a variety of forms.<sup>6</sup> Some are temporal. Alexander Graham Bell got some benefit from his invention of the telephone, but he doesn't anymore. The companies he founded continue to make money from the telephone, but they also have competitors who make money from the invention as well, and users of the telephone benefit in countless ways for which they do not fully pay. Other spillovers are geographic, in several different senses. An innovation contributes to a local economy, employing people who spend their money locally. Geographic spillovers may also occur in a very different sense: Population density is strongly positively related to local innovation—people are more inventive when they are around other inventive people.<sup>7</sup> Ideas may travel further or more

---

Council of Econ. Advisors, Supporting Research and Development to Promote Economic Growth: The Federal Government's Role 7 tbl.2 (1995).

The sources cited in the table are listed in the order presented: M. Ishaq Nadiri, Innovations and Technological Spillovers (Nat'l Bureau of Econ. Research, Working Paper No. 4423, 1993), available at <http://papers.nber.org/papers/W4423> (on file with the *Columbia Law Review*); Edwin Mansfield et al., Social and Private Rates of Return from Industrial Innovations, 91 Q.J. Econ. 221 (1977); Nestor E. Terleckyj, Nat'l Planning Ass'n, Effects of R&D on the Productivity Growth of Industries: An Exploratory Study (1974); Leo Sveikauskas, Technology Inputs and Multifactor Productivity Growth, 63 Rev. Econ. & Stat. 275 (1981); Akira Goto & Kazuyuki Suzuki, R&D Capital, Rate of Return on R&D Investment and Spillover of R&D in Japanese Manufacturing Industries, 71 Rev. Econ. & Stat. 555 (1989); Jeffrey I. Bernstein & M. Ishaq Nadiri, Interindustry R&D Spillovers, Rates of Return, and Production in High-Tech Industries, 78 Am. Econ. Rev. (Papers & Proc.) 429 (1988); Frederick M. Scherer, Using Linked Patent and R&D Data to Measure Interindustry Technology Flows, in R&D, Patents, and Productivity 417 (Zvi Griliches ed., 1984); Jeffrey Bernstein & M. Ishaq Nadiri, Product Demand, Cost of Production, Spillovers, and the Social Rate of Return to R&D (Nat'l Bureau of Econ. Research, Working Paper No. 3625, 1991), available at <http://www.nber.org/papers/W3625> (on file with the *Columbia Law Review*); see also Zvi Griliches, The Search for R&D Spillovers, 94 Scandinavian J. Econ. S29, S43 (1992) ("In spite of [many] difficulties, there has been a significant number of reasonably well done studies all pointing in the same direction: R&D spillovers are present, their magnitude may be quite large, and social rates of return remain significantly above private rates.").

For additional discussion of the gap between private and social benefits, see generally, e.g., Timothy F. Bresnahan, Measuring the Spillovers from Technical Advance: Mainframe Computers in Financial Services, 76 Am. Econ. Rev. 742, 753 (1986) (finding large spillover effect from adoption of computer in financial services sector of economy); Paul A. David, Bronwyn H. Hall & Andrew A. Toole, Is Public R&D a Complement or Substitute for Private R&D? A Review of the Econometric Evidence, 29 Res. Pol'y 497 (2000) (surveying research on social rate of return to R&D); Bronwyn H. Hall, The Private and Social Returns to Research and Development: What Have We Learned?, in Technology, R&D, and the Economy 140 (Bruce L.R. Smith & Claude E. Barfield eds., 1996) (same); Pierre Mohnen, R&D Externalities and Productivity Growth, 18 STI Rev. 39 (1996) (same).

6. Frischmann has explained how copyright promotes spillovers. See Brett M. Frischmann, Evaluating the Demsetzian Trend in Copyright Law, 2 Rev. L. & Econ. (forthcoming 2006) (manuscript at 12–13, on file with the *Columbia Law Review*) [hereinafter Frischmann, Demsetzian Trend]; see also *infra* Part III.A.

7. See Gerald Carlino et al., Matching and Learning in Cities: Urban Density and the Rate of Invention 17–23 (Fed. Reserve Bank of Phila., Working Paper No. 04-16/R, 2005), available at <http://ssrn.com/abstract=598127> (on file with the *Columbia Law Review*). The authors offer several reasons for this result—not just scale economies and the availability of

quickly than they would if confined to market transactions.<sup>8</sup> Other spillovers are interindustry: Work done in one field, such as defense or space science, may have benefits in seemingly unrelated fields such as materials science. The inventor may have no interest in or even awareness of the benefits her idea has in these unrelated fields. Still other spillovers occur within an industry; as the Bricklin example illustrates, it is impossible to exclude competitors or other researchers entirely from the intellectual benefits of an innovation. If nothing else, people will learn that something is possible from the fact that someone else has done it.

Spillovers may also be realized by consumers and community members even where parties have transacted over the distribution of surplus. During the period Bell held patents on the telephone, he did not recognize, much less capture, the full benefits of his invention. People paid him a certain amount for telephones, but as others bought phones, the value of the phones to the earlier purchasers went up, and Bell did not capture any of that new value.<sup>9</sup> Nor did he capture the benefits to those who saved money or lives because their neighbors could call the police to report a crime, or any of countless other benefits the telephone provided to society. Of course, not all of the difference between producer and social surplus in a transaction should be characterized as a spillover. To the extent that the parties transact and recognize the sharing of surplus between them, then the benefits are not really external to the transaction.<sup>10</sup> In the examples we give, however, the particular benefits are not captured in the transaction because they either constitute unanticipated consumer surplus (from network effects, new uses, etc.) or benefits realized by third parties not party to any transaction.

---

human capital, but the ability to be selective in identifying team members, which leads to more productive working environments. *Id.* at 7–9.

8. For a discussion of how inventions disseminated before the modern era, see generally Petra Moser, *How Do Patent Laws Influence Innovation? Evidence from Nineteenth-Century World Fairs* (Nat'l Bureau of Econ. Research, Working Paper No. W9909, 2003), available at [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=435483](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=435483) (on file with the *Columbia Law Review*).

9. For further discussion and analysis of “network effects”—markets in which the value the consumer places on the good increases as others use the good—see Nicholas Economides, *The Economics of Networks*, 14 *Int'l J. Indus. Org.* 673, 678–91 (1996) (discussing salient economic features of network externalities); Michael L. Katz & Carl Shapiro, *Network Externalities, Competition, and Compatibility*, 75 *Am. Econ. Rev.* 424, 424–25 (1985) (discussing effects of networks on competition and market structure as well as issue of compatible products); Michael L. Katz & Carl Shapiro, *Systems Competition and Network Effects*, *J. Econ. Persp.*, Spring 1994, at 93, 96–97 (discussing communications networks). See generally Mark A. Lemley & David McGowan, *Legal Implications of Network Economic Effects*, 86 *Cal. L. Rev.* 479 (1998) (discussing ways in which network effects should affect various legal fields).

10. See John F. Duffy, *Intellectual Property Isolationism and the Average Cost Thesis*, 83 *Tex. L. Rev.* 1077, 1081–85 (2005) (highlighting difference between consumer surplus and externalities if parties transact and recognize surplus presence).

## B. *The Law and Economics Account of Spillovers*

1. *Ignoring Pecuniary Externalities.* — Spillovers are a species of externality, which itself has been a contested concept in economics for many years.<sup>11</sup> For our purposes, positive (or negative) externalities are benefits (costs) realized by one person as a result of another person's activity without payment (compensation). Externalities generally are not fully factored into a person's decision to engage in the activity. Throughout this paper, we refer to spillovers as externalities and, at times, more generically as the difference between private and social returns. Yet not all spillovers result in a wedge between private and social returns. Those that do so are referred to by economists as "real" or "technological" externalities, while those that fail to do so are deemed "pecuniary" externalities. The former type of externality affects net social welfare while the latter type only affects distribution (e.g., the division of surplus between producers and consumers).<sup>12</sup>

Technological externalities are direct benefits (or costs) realized by third parties—agents who are not participating in the relevant market and thus have not transacted with the provider of the benefits or costs. Think here of the diffuse social benefits of education or the social costs of environmental pollution. Economists have long recognized that their existence can cause social welfare to diverge from private welfare. We discuss them below.

Pecuniary externalities, by contrast, constitute wealth transfers among private parties. Pecuniary externalities generally affect participants in a market, or parties to a transaction, by operating on the price mechanism. The most common sort involve differences between what a consumer is willing to pay for a product and what the consumer is actually required to pay in the market. I may be willing to pay \$100 to buy the first season of *Battlestar Galactica* on DVD;<sup>13</sup> if the producer only charges me \$38.99,<sup>14</sup> the remaining \$61.01 is a windfall to me. This "consumer surplus" is *external* to the producer's activities and decisionmaking—the producer cares only that I am willing to pay at least \$38.99, not how much more I value it—but it is a pecuniary externality because the DVD is voluntarily exchanged.<sup>15</sup>

---

11. See Andreas A. Papandreou, *Externality and Institutions* 13–68 (1994) (providing detailed historical account of term); Harold Demsetz, *Toward a Theory of Property Rights*, 57 *Am. Econ. Rev. (Papers & Proc.)* 347, 348 (1967) [hereinafter Demsetz, *Theory*] ("Externality is an ambiguous concept.").

12. While this seems like a nice dichotomy, identifying and distinguishing between these types of externalities is not straightforward. Moreover, as we discuss below, spillovers that seem to be pecuniary externalities may turn out to be technological externalities upon closer inspection.

13. The new series, of course, not the cheesy one from the 1980s.

14. See [www.amazon.com](http://www.amazon.com) (last searched Feb. 20, 2006).

15. Some economists would say that the consumer surplus is not an externality at all because the seller is aware of and takes into account the excess value realized by consumers. But that is incorrect. Sellers take buyers' willingness to pay as a signal for

Pecuniary externalities aren't limited to buyer-seller transactions. They may encompass certain types of market substitution with offsetting benefits and costs. As Posner notes,

[c]ompetition is a rich source of "pecuniary" externalities, as distinct from "technological" externalities—that is, of wealth transfers from, as distinct from cost impositions on, unconsenting parties. Suppose A opens a gas station opposite B's gas station and as a result siphons revenues from B. Since B's loss is A's gain, there is no diminution in overall wealth and hence no social cost, even though B is harmed by A's competition and thus incurs a private cost.<sup>16</sup>

With the notable exception of Pigou,<sup>17</sup> economists don't much care about pecuniary externalities, reasoning that wealth transfers "within" the market—that is, externalities mediated by the price mechanism—result in offsetting private costs and benefits. Economists are similarly unconcerned with wealth transfers "within" a transaction. If the parties were both willing participants in a transaction, it was because the buyer valued the product at least as much as the seller, creating a surplus. Society—or at least economics—has no particular interest in how that surplus is divided.<sup>18</sup>

Some have suggested as a result that we can safely ignore most innovation spillovers, since they are just pecuniary externalities.<sup>19</sup> Just create property rights, the argument goes, and the owners will sell to those who value the use more. In fact, however, distinguishing between real and pecuniary externalities is very difficult in the innovation context. As we discuss in the next section, property rights in innovation are justified by arguments about internalizing both "real" externalities *and* pecuniary externalities, on the theory that uncaptured innovation-generated value in the hands (or minds) of innovation users is said to have negative external effects on the dynamic incentives of third-party innovators.

It turns out that once we are willing to entertain the idea that the allocation of rights and thus wealth may have dynamic external effects, distinguishing technological and pecuniary externalities is much more difficult than one would think. Essentially, what look like pecuniary ex-

demand, and that signal is what guides sellers' decisions. Generally, sellers have no reason to take into account consumer surplus, absent a realistic opportunity to price discriminate.

16. Richard A. Posner, *Economic Analysis of Law* 7 (6th ed. 2003). Some would challenge even this claim. See Michael Abramowicz, *An Industrial Organization Approach to Copyright Law*, 46 *Wm. & Mary L. Rev.* 33, 80–88 (2004).

17. Arthur Cecil Pigou, *The Economics of Welfare* (Transaction Publishers 2002) (1920).

18. For more discussion, see Duffy, *supra* note 10, at 1081–82, and sources cited therein.

19. See *id.* at 1082, and sources cited therein; David D. Haddock, *Irrelevant Internalities, Irrelevant Externalities, and Irrelevant Anxieties* 4 (Nw. Univ. Sch. of Law, Law & Econ. Research Paper Series, Research Paper No. 03-16, 2003), available at [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=437221](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=437221) (on file with the *Columbia Law Review*) ("Externalities are everywhere but usually economically meaningless.").

ternalities from a static perspective look more like technological externalities from a dynamic perspective in situations where wealth transfers affect incentive structures, behavior, and resource allocation. We consider which externalities really matter in Part II.A.

2. *Cleaning Up Third-Party Spillovers.* — On the standard law and economics account of property,<sup>20</sup> technological externalities—spillovers that affect third parties—cannot be ignored. Rather, they are a bad thing that ought to be minimized if not eliminated. At least since Harold Demsetz’s seminal article, *Toward a Theory of Property Rights*,<sup>21</sup> law and economics

20. We focus on this particular subdiscipline because it has been very influential in affecting the rhetoric and rationale for absolute property rights—in the field of IP and other fields as well. See Mark A. Lemley, *Property, Intellectual Property, and Free Riding*, 83 *Tex. L. Rev.* 1031, 1033–46 (2005) [hereinafter Lemley, *Free Riding*] (noting rise in treatment of intellectual property in terms of real property rights). It is important to recognize that the particular views held by this subdiscipline are not universally held by economists. Economists who study innovation, industrial research and development, endogenous growth, and many other subdisciplines view spillovers as beneficial and at times worth encouraging. Government funding of basic research, for example, is generally justified on the grounds that spillovers will occur: the more, the merrier. See, e.g., Council of Econ. Advisors, *supra* note 5, at 15 (arguing that federal government should “fill the gap between the private level of R&D investment and the level and types of R&D investment that are best for the nation”); Dan Galai & Zvi Wiener, *Government Support of Investment Projects in the Private Sector: A Microeconomic Approach*, *Fin. Mgmt.*, Autumn 2003, at 33, 35–36 (discussing R&D spillover effect). The seminal works making the market failure argument for government support of basic research are Richard R. Nelson, *The Simple Economics of Basic Scientific Research*, 67 *J. Pol. Econ.* 297 (1959), and 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 (Richard Nelson ed., 1962) [hereinafter Arrow, *Economic Welfare*].

21. Demsetz, *Theory*, *supra* note 11. Demsetz famously observed that private property rights emerge and evolve as the value of things increases, and technology and new markets develop to capture that value. His article has been extremely influential both in explaining the emergence and shape of property rights in various societies and in providing a platform for making normative arguments about extending property rights. Demsetz’s work can be read fairly to suggest both a *descriptive thesis*—that private property rights in fact emerge to enable the internalization of externalities as the value of resources increases, and technologies and markets emerge to make internalization less costly (more beneficial)—and a *normative thesis*—that the emergence of private property rights to enable the internalization of externalities is desirable (in an economic framework, social welfare maximizing). See Frischmann, *Demsetzian Trend*, *supra* note 6, *passim*. As a matter of prediction, Demsetz’s descriptive thesis hardly seems controversial—as the value of resources increases, private property will emerge, adapt, or evolve to better enable appropriation of value. Demsetz does not specify the process by which property rights evolve. See Thomas W. Merrill, *Introduction: The Demsetz Thesis and the Evolution of Property Rights*, 31 *J. Legal Stud.* 331, 333 (2002) (outlining questions unresolved by Demsetz thesis). It is as if the dynamic being described is a natural process, as if the appropriation of value via private property rights is a natural evolutionary endpoint. See Stuart Banner, *Transitions Between Property Regimes*, 31 *J. Legal Stud.* 359, 360 (2002); see also Saul Levmore, *Two Stories About the Evolution of Property Rights*, 31 *J. Legal Stud.* 421, 422–23 (2002) (describing progression from “‘commons’ with open access” to “single private property owners with virtually full control, who normally close access or charge for it” as “normal evolution”); Merrill, *supra*, at 332 (noting Demsetz’s hypothesis

scholars have thought about private property rights as a means of encouraging both efficient allocation of private investment into the creation of resources and efficient management of resources.<sup>22</sup> On a strong version of Demsetz's normative thesis,<sup>23</sup> private property becomes an effective means of achieving these ends when it leads to the complete internalization of externalities.

It is important to emphasize that it is the *social* costs and *social* benefits of internalization that need to be weighed before creating or extending private property rights.<sup>24</sup> The basic idea behind "internalizing externalities" is that if property owners are both fully encumbered with potential third-party costs and entitled to completely appropriate potential third-party benefits, their interests will align with the interests of society, and they will make efficient (social welfare-maximizing) decisions.<sup>25</sup> For example, if land owners are forced to internalize the costs polluting the air or water might have on neighbors, they will pollute efficiently—that is, only to the extent that the total social benefits of doing so exceed

---

on when property rights emerge); cf. Ben Depoorter, *The Several Lives of Mickey Mouse: The Expanding Boundaries of Intellectual Property Law*, 9 Va. J.L. & Tech. 4, ¶ 48 (2004), at <http://www.vjolt.net/archives/php?issue=18> (on file with the *Columbia Law Review*) ("[T]here actually is a degree of determinacy in the evolution of the laws of intellectual property."). Of course, it is not.

22. As Frischmann notes:

Private property rights economize on both the costs of exclusion and transaction and thereby facilitate appropriation of surplus generated through the use of private resources. It is the improved capacity to appropriate surplus—whether through one's own private exclusive use or by transferring one's rights to another—that improves incentives to invest in the production, development, and management of property.

Frischmann, *Demsetzian Trend*, supra note 6 (manuscript at 6–7) (footnote omitted).

23. See supra note 21. We are not convinced that Demsetz necessarily meant to make a strong normative argument at all. Still, much like the way in which interpretations of Ronald Coase's scholarship have taken his original insights in unintended directions, Demsetz's property theory has likewise been extended to support normative arguments for increased propertization and privatization of valuable resources. Our focus is on the normative theory drawn from Demsetz's original work; we refer to this normative theory and arguments derived from it as *Demsetzian*, while acknowledging that he might not himself subscribe to the arguments made in his name.

24. Cf. Demsetz, *Theory*, supra note 11, at 350 ("[P]roperty rights develop to internalize externalities when the gains of internalization become larger than the cost of internalization."); Duffy, supra note 10, at 1080–85 (recognizing the need to balance costs and benefits of property rights); Gideon Parchomovsky & Peter Siegelman, *Selling Mayberry: Communities and Individuals in Law and Economics*, 92 Cal. L. Rev. 77, 79–80 (2004) (noting that economists identify "high transaction costs as the key barrier to the efficient internalization of externalities such as pollution"). See generally Symposium, *The Evolution of Property Rights*, 31 J. Legal Stud. 331 (2002) (presenting series of articles analyzing Demsetz's theory on evolution of property rights from various viewpoints).

25. According to Demsetz, "[e]very cost and benefit associated with social interdependencies is a potential externality." Demsetz, *Theory*, supra note 11, at 348. Actual externalities exist where benefits or costs are not taken into account by interacting parties because "[t]he cost of . . . transact[ing] in the rights between the parties (internalization) must exceed the gains from internalization." *Id.*

the total social costs. They must also internalize benefits in order to have the proper incentive to invest in maintaining and improving their property. According to the Demsetzian theory, internalization is the silver bullet that magically aligns private and social welfare.

The Demsetzian view is more or less the same in the context of innovation. If an inventor cannot capture the full social benefit of her innovation, the argument goes, she will not have enough incentive to engage in the research and development that will produce that innovation. If there are spillovers from innovation,<sup>26</sup> they must be interfering with incentives to innovate, and we should find them and stamp them out. Further, and of more direct relevance, only with complete internalization will an inventor be able to efficiently manage an innovation after it is created.

This latter “management” argument for full internalization is the basis of the prospect theory of IP pioneered by Ed Kitch.<sup>27</sup> On this theory, granting exclusive rights over a particular invention will lead people to manage, develop, and commercialize that invention efficiently, because they will capture the full social benefits of doing so. Here again spillovers—both technological and (at least some) pecuniary externalities<sup>28</sup>—are a problem to be solved by internalization. On the prospect rationale, if competitors can benefit from an invention without transacting with the owner of that invention for permission, the owner won’t have the proper incentives to manage, market, and sell products implementing the invention.<sup>29</sup>

Demsetz and Kitch draw support from the analytic framework of Ronald Coase.<sup>30</sup> Coase argued that in a world of no transaction costs, all

26. See Baumol, *supra* note 5, at 5–6 (noting argument that spillovers impede innovation). As noted in the previous section, innovation spillovers include both technological and pecuniary externalities. While pecuniary externalities, as mere wealth transfers, are ordinarily ignored by economists, they play an important role in innovation policy because of their potential dynamic effects.

27. See Edmund W. Kitch, *The Nature and Function of the Patent System*, 20 *J.L. & Econ.* 265, 276 (1977); F. Scott Kieff, *Property Rights and Property Rules for Commercializing Inventions*, 85 *Minn. L. Rev.* 697, 717 (2001); see also Shubha Ghosh, *Patents and the Regulatory State: Rethinking the Patent Bargain Metaphor After Eldred*, 19 *Berkeley Tech. L.J.* 1315, 1353–57 (2004) [hereinafter Ghosh, *Rethinking*] (noting that prospect and commercialization theories derive from theoretical work of Demsetz).

28. See *supra* text accompanying notes 18–19.

29. It is worth noting that even some economists not normally associated with prospect theory occasionally lapse into the assumption that spillovers need to be internalized. See, e.g., James Bessen & Michael J. Meurer, *Lessons for Patent Policy from Empirical Research on Patent Litigation*, 9 *Lewis & Clark L. Rev.* 1, 6 (2005). As Paul Goldstein argued with respect to copyright: “The logic of property rights dictates their extension into every corner in which people derive enjoyment and value from literary and artistic works. To stop short of these ends would deprive producers of the signals of consumer preference that trigger and direct their investments.” Paul Goldstein, *Copyright’s Highway: From Gutenberg to the Celestial Jukebox* 178–79 (1994).

30. R.H. Coase, *The Problem of Social Cost*, 3 *J.L. & Econ.* 1, 2–6 (1960) [hereinafter Coase, *Social Cost*].

we need are “well-defined” property rights.<sup>31</sup> It does not even matter to whom property rights are assigned (putting aside “questions of equity”) because everyone who might be affected by the use of the resource to which the property right applies will bargain to reallocate rights in a manner that maximizes social welfare. Of course, “well-defined” does not really say anything about the strength, scope, duration, or other important dimensions of those property rights. But where the lines are drawn doesn’t really matter in a world of zero transaction costs. All you need is a baseline from which to bargain to the optimal arrangement.<sup>32</sup>

But as Coase recognized, we live in a world rife with transaction costs. He mainly intended to emphasize the importance of taking into account transaction costs when comparing institutional solutions to perceived market failures. Property rights can be a useful institution for reducing transaction costs where such rights delegate decisionmaking authority to the entity with the closest connection to and best information regarding the resource. Those in the Demsetzian property rights school presume that the inventor is the best informed party (or the best information gatherer)<sup>33</sup> and thus the optimal manager or steward, and alternatively fall back on the assumption that if the inventor isn’t the best manager for a particular use, she will license the innovation to someone who is.

The obvious implication of the property rights theory is that spillovers are bad, since they drive a wedge between private and social value and prevent the perfectly informed inventor from making optimal decisions. From the supply side, spillovers are uncaptured benefits that could be captured to increase incentives to invest, and from the demand side, spillovers reflect unobserved, lost signals of consumer demand that fail to guide investment and management decisions. For a property

---

31. Property rights might be unnecessary altogether in a world of zero transaction costs because contract law would be adequate. See Matthew J. Sag, *Beyond Abstraction: The Law and Economics of Copyright Scope and Doctrinal Efficiency*, 81 *Tul. L. Rev.* (forthcoming 2006) (manuscript at 18, on file with the *Columbia Law Review*) [hereinafter Sag, *Beyond Abstraction*] (“In a Coasian world we could replace copyright with contract entirely, or we could have neither and rely on the government to reward each author with the exact social value of her creation.”). But see Paul J. Heald, *A Transaction Costs Theory of Patent Law*, 66 *Ohio St. L.J.* 473, 481 (2005) (noting that contracting involves transaction costs).

32. Coase, *Social Cost*, *supra* note 30, at 19 (“[I]f market transactions were costless, all that matters (questions of equity apart) is that the rights of the various parties should be well-defined and the results of legal actions easy to forecast.”).

33. See Henry E. Smith, *Property and Property Rules*, 79 *N.Y.U. L. Rev.* 1719, 1754–56 (2004). Henry Smith argues that property rights (coupling the right to exclude with a property rule) involve a delegation to owners to gather and act upon information regarding the uses of a resource. He explains the pervasiveness of property rights in the law as a result of a strong preference for decentralized information gathering and decisionmaking by property owners. Whatever the merits of Smith’s argument for real property, it doesn’t work for IP. IP rights, as an artificial deviation from competition, centralize rather than decentralize information gathering and decisionmaking. Centralization is in fact at the heart of the Demsetz-Kitch approach.

rights theorist, we could solve a great number of problems if only we could eliminate spillovers.

### C. *The Benefits of Spillovers*

In fact, however, a wealth of economic evidence teaches us that spillovers are good for society. There is no question that inventions create significant social benefits beyond those captured in a market transaction. Statistical evidence repeatedly demonstrates that innovators capture only a small proportion of the social value of their inventions.<sup>34</sup> Some of these spillovers accrue to passive consumers in the form of consumer surplus.<sup>35</sup> But more importantly, spillovers also benefit third parties, including competitors and potential competitors. As employees leave an innovative company, or as other companies learn from their dealings with that company or from competing with it, they acquire knowledge about innovative products and processes that they can put to use elsewhere.<sup>36</sup> The public as a whole also benefits.<sup>37</sup>

Far from interfering with incentives, these spillovers actually drive further innovation. Industries with significant spillovers generally experience more and faster innovation than industries with fewer spillovers.<sup>38</sup> Similarly, cities with significant spillovers experience more innovation than smaller areas that can limit employee mobility.<sup>39</sup> Contrary to what a

---

34. See *supra* note 5 (collecting that evidence).

35. We recognize that not all consumer surplus constitutes a spillover. Still, as we discussed earlier, there are instances where benefits that accrue to consumers are not taken into account in transactions between producers and consumers. See *supra* note 5 and accompanying text.

36. See, e.g., Holger Görg & Alexander Hijzen, *Multinational Enterprises and Spillovers*, in *Globalisation and Productivity Growth* 92, 92–93 (Holger Görg et al. eds., 2005) (noting ways in which spillovers to competitors occur); Keller & Yeaple, *supra* note 5, at 26–28 (documenting their significance).

37. See Baumol, *supra* note 5, at 121–23.

38. See, e.g., Dietmar Harhoff, *R&D Spillovers, Technological Proximity, and Productivity Growth—Evidence from German Panel Data*, 52 *Schmalenbach Bus. Rev.* 238, 258 (2000) (“High-technology firms react more sensitively to spillovers in terms of their R&D spending, and their direct marginal productivity gain from spillovers (in excess to the effect from enhanced R&D spending) is considerably larger than the respective gain for less technology-oriented firms.”). Indeed, the positive relationship is so strong that some economists use spillovers as a measure of innovation! See Tobias Schmidt, *An Empirical Analysis of the Effects of Patent and Secrecy on Knowledge Spillovers 1* (Ctr. for Eur. Econ. Res., Discussion Paper No. 06-048, 2006), available at <ftp://ftp.zew.de/pub/zew-docs/dp/dp06048.pdf> (on file with the *Columbia Law Review*). For detailed economic discussion of the role of spillovers on innovation, see generally Claude d’Aspremont & Alexis Jacquemin, *Cooperative and Noncooperative R&D in an Oligopoly with Spillovers*, 78 *Am. Econ. Rev.* 1133 (1988); Kotaro Suzumura, *Cooperative and Noncooperative R&D in Duopoly with Spillovers*, 82 *Am. Econ. Rev.* 1307 (1992); Irene Henriques, *Comment, Cooperative and Noncooperative R&D in Duopoly with Spillovers*, 80 *Am. Econ. Rev.* 638 (1990).

39. See Carlino et al., *supra* note 7, at 31 (providing empirical results indicating that “the number of inventions per person is about 20 percent greater in [a metropolitan area] with a local economy that is twice as dense as another [metropolitan area]”). See generally

reader of property rights theory might predict, Harhoff finds empirical evidence that firms in high-technology industries (the most innovation-intensive ones) are likely to increase rather than decrease their R&D investment in the face of significant intra-industry spillovers.<sup>40</sup> Acs et al. argue that this is because the spillovers are creating opportunities to be exploited by entrepreneurs.<sup>41</sup> But these entrepreneurs aren't engaging in incentive-draining free riding; rather, they are part of a virtuous circle because they are in turn creating new knowledge spillovers that support still more entrepreneurial activity. This is consistent with other work finding that spillovers may *increase* incentives for private investment in R&D due to complementarities in spillover processes.<sup>42</sup> For example, investments in R&D may increase a firm's capacity to absorb spillovers from competitors and/or other industries altogether. Appreciating the potential benefits of spillovers may depend upon a firm's (technical) capacity to recognize, understand, and utilize an innovation.<sup>43</sup> Firms have an incentive to develop "absorptive capacity" for inevitable spillovers by investing in R&D.<sup>44</sup>

---

Annalee Saxenian, *Regional Advantage: Culture and Competition in Silicon Valley and Route 128* (1994) (discussing value of geographic proximity in producing spillovers that drive innovation).

40. Harhoff, *supra* note 38, at 258. By contrast, Harhoff finds that in low-tech industries, spillovers tend to have an R&D-dampening effect more consistent with the Demsetz hypothesis. *Id.*; see also Jeffrey I. Bernstein & M. Ishaq Nadiri, *Research and Development and Intra-Industry Spillovers: An Empirical Application of Dynamic Duality*, 56 *Rev. Econ. Stud.* 249, 257–58 (1989) (providing empirical data demonstrating that high intra-industry spillover tends to reduce firms' R&D investments in low-technology industries).

41. Zoltan J. Acs et al., *The Knowledge Spillover Theory of Entrepreneurship* 23 (Ctr. for Econ. Policy Research, Discussion Paper No. 5326, 2005), available at [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=873614](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=873614) (on file with the *Columbia Law Review*).

42. See, e.g., Wesley M. Cohen & Daniel A. Levinthal, *Innovation and Learning: The Two Faces of R&D*, 99 *Econ. J.* 569, 569–70 (1989) (arguing that investment in R&D "develops the firm's ability to identify, assimilate, and exploit knowledge from the environment," which indicates that greater knowledge in the environment creates greater incentives for R&D investment); Georg von Graevenitz, *Spillovers Reconsidered: Analysing Economic Welfare Under Complementarities in R&D* 16 (*Governance & the Efficiency of Econ. Sys.*, Discussion Paper No. 29, 2004), available at <http://www.gesy.uni-mannheim.de/dipa/29.pdf> (on file with the *Columbia Law Review*) (presenting spillover R&D model indicating that there are complementarities in private actors' R&D efforts through spillover processes).

43. R&D investments not only yield research results but also educate, train, and enable researchers to absorb spillovers from competitors. Economists refer to this as absorptive capacity. The basic idea is tied to the fact that engaging in the research process builds human capital and educates—it does more than just yield information or research outputs. See Brett M. Frischmann, *The Pull of Patents 9* (*ExpressO Preprint Series*, Working Paper No. 700, 2005), available at <http://law.bepress.com/expresso/eps/700> (on file with the *Columbia Law Review*).

44. Harhoff, *supra* note 38, at 239. Firms also have an incentive to encourage rivals to similarly invest in R&D because of the spillovers those investments might generate.

Further, the absorptive capacity investments don't simply represent market stealing. Rather, they are complementary to the initial firm's investments,<sup>45</sup> and therefore drive new innovation and improve productivity. Harhoff and Klette both find that spillovers enhance firm productivity in innovation-intensive industries—the same industries that invest in R&D to absorb spillovers.<sup>46</sup> The computer industry is an obvious example. Both Annalee Saxenian and Ron Gilson have shown that spillovers drove innovation in that industry: Silicon Valley thrived while Boston's Route 128 withered in the 1980s and 1990s in significant part because employees and knowledge moved freely to new companies in Silicon Valley, but not in Boston.<sup>47</sup> Spillovers also drive innovation in other industries in which the original inventor does not even participate.<sup>48</sup> And university-originated knowledge spillovers drive significant economic

---

45. Adam B. Jaffe, *Demand and Supply Influences in R&D Intensity and Productivity Growth*, 70 *Rev. Econ. & Stat.* 431, 437 (1988).

46. Harhoff, *supra* note 38, at 258; T.J. Klette, *R&D, Spillovers and Performance Among Heterogeneous Firms: An Empirical Study Using Microdata* (Statistics Nor., Discussion Paper No. 133, 1994).

47. See Saxenian, *supra* note 39, at 161–68; Ronald J. Gilson, *The Legal Infrastructure of High Technology Industrial Districts: Silicon Valley, Route 128, and Covenants Not to Compete*, 74 *N.Y.U. L. Rev.* 575, 577–78 (1999). Those spillovers are local in nature, because further evidence suggests that the benefits from spillovers are concentrated in a local area rather than being widely dispersed, since employees often depart for jobs in the same region, and many suppliers are local as well. For empirical support for the Saxenian-Gilson hypothesis, see Bruce Fallick et al., *Job-Hopping in Silicon Valley: Some Evidence Concerning the Microfoundations of a High-Technology Cluster*, 88 *Rev. Econ. & Stat.* 472, 475–81 (2006). For economic literature discussing locality spillovers, see, e.g., David B. Audretsch, *Agglomeration and the Location of Innovative Activity*, 14 *Oxford Rev. Econ. Pol'y* 18, 26 (1998) (discussing spillovers and innovation in global markets); David B. Audretsch & Maryann P. Feldman, *R&D Spillovers and the Geography of Innovation and Production*, 86 *Am. Econ. Rev.* 630, 639 (1996) (presenting “evidence that industries in which knowledge spillovers are more prevalent . . . have a greater propensity for innovative activity to cluster”); Rui Baptista, *Clusters, Innovation, and Growth: A Survey of the Literature*, in *The Dynamics of Industrial Clustering: International Comparisons in Computing and Biotechnology* 13, 14 (G.M. Peter Swann, Martha Prevezer & David Stout eds., Oxford Univ. Press 1998) (hypothesizing that if there are “spillovers to be gained from a strong core,” there will be clustering within industry if “companies are located close to the core”); Rui Baptista, *Do Innovations Diffuse Faster Within Geographical Clusters?*, 18 *Int'l J. Indus. Org.* 515, 516 (2000) (arguing that “clustering of innovators” causes “regional differences in diffusion rates”); Michael Storper & Bennett Harrison, *Flexibility, Hierarchy and Regional Development: The Changing Structure of Industrial Production Systems and Their Forms of Governance in the 1990s*, 20 *Res. Pol'y* 407, 411–14 (1991) (noting “wide range of types of power and hierarchy in the governance of a production system”); Michael Storper, *Regional Technology Coalitions: An Essential Dimension of National Technology Policy*, 24 *Res. Pol'y* 895, 900 (1995) (“[S]pillovers are economic, in that the pecuniary benefits . . . will ‘leak’ from firm to firm within a technological space due to the ways that certain innovations enhance the capacities of other firms.”).

48. See, e.g., Jochen Streb et al., *Technological and Geographical Knowledge Spillover in the German Empire 1877–1918*, 59 *Econ. Hist. Rev.* 347 (2006).

growth,<sup>49</sup> notwithstanding the theory that university ideas will languish unless universities have patent rights over them.

The positive effects of spillovers present a puzzle for the property rights theory of IP. If we have systematically failed to internalize the benefits of innovation, why does society seem to be benefiting rather than suffering from this failure? We seek to answer that question in Part II.

## II. UNDERSTANDING THE BENEFITS OF SPILLOVERS

In this Essay, we offer a way to solve the puzzle of positive spillovers by building on three core insights, one focused on which spillovers matter, one focused on the supply side (the costs of spillovers to innovation incentives),<sup>50</sup> and the last on the demand side (the social benefits of spillovers).<sup>51</sup>

### A. *What Spillovers Matter? Spillovers Within and Outside Transactions*

The economist's decision to ignore pecuniary externalities in setting policy makes a certain amount of sense in circumstances where parties are making and exchanging private goods.<sup>52</sup> Private goods are consumed rivalrously and thus are naturally scarce.<sup>53</sup> Real and personal property

49. See David B. Audretsch et al., *The Knowledge Spillover Theory of Entrepreneurship and Technological Diffusion*, in *University Entrepreneurship and Technology Transfer: Process, Design, and Intellectual Property* 69, 79 (Gary D. Libecap ed., 2005).

50. See Lemley, *Free Riding*, supra note 20, at 1068; see also Mark A. Lemley, *What's Different About Intellectual Property?*, 83 *Tex. L. Rev.* 1097, 1099–102 (2005) [hereinafter Lemley, *What's Different?*] (further elaborating theory and answering objections).

51. Frischmann, *Economic Theory*, supra note 3, at 968. Theories of IP are myopically supply sided. The reward, prospect, and commercialization theories emphasize IP-enabled exclusivity as the relevant means for fixing a supply-side problem—essentially, the undersupply of private investment into the production of patentable or copyrightable subject matter or into the development and commercialization of such subject matter that would occur in the absence of IP-enabled exclusivity. The theories differ largely in terms of where in the supply chain IP-enabled exclusivity is needed and of the degree of control/exclusivity needed to attract private investment. Surprisingly, these theories largely ignore demand-side considerations and take for granted that consumers' willingness to pay will accurately signal societal demand. See Brett M. Frischmann, *Commercializing University Research Systems in Economic Perspective: A View from the Demand Side*, in *University Entrepreneurship and Technology Transfer: Process, Design, and Intellectual Property*, supra note 49, at 155, 176 [hereinafter Frischmann, *Commercializing University Research Systems*] (stating that reliance on market leads to distortion in measuring demand). As Frischmann's prior work explores, and as we discuss below, spillovers frequently arise in situations where beneficiaries' willingness to pay understates societal demand.

52. Portions of this paragraph and the next are adapted from Lemley, *What's Different?*, supra note 50, at 1100–01.

53. If A consumes a private good—say an apple—then B's consumption opportunities diminish because the good is depleted. While it is true that even apples can be used productively in a nonrivalrous fashion—for example, when a person views an apple and produces a painting of it, the apple is not depleted—the most prevalent uses of an apple involve rivalrous consumption. In addition, private goods generally are rivalrously *possessed*, meaning that A's possession of an apple precludes B from possessing the apple

rights systems generally establish strong ownership rights in the first possessor of such goods. This facilitates market transactions. It is widely accepted that social welfare is maximized when a rivalrous good is consumed by the person that values it most and that the market mechanism is generally the most efficient means for rationing such goods and for allocating resources needed to produce them.<sup>54</sup>

This market system for real and personal property rights works because the nature and scope of those rights are, in the main, quite clear.<sup>55</sup> We can all find out what the boundaries of real property are, either by looking at physical fences or by going down to the county recorder's office and looking up the boundaries. We also have a good idea what the legal rules are with respect to property—physical intrusion is generally forbidden, and other kinds of intrusion generally aren't. The physical possession of property means that when someone wants land or a good belonging to someone else, he knows he needs to transact to get it, and he can quite easily find out with whom he must deal. When those transactions occur, economists feel safe in ignoring the distribution of costs and benefits to the parties to the transaction itself.

The transactions that create pecuniary externalities presuppose a baseline set of property rights. For land and goods, we take those background property rights for granted, both because the rules have been with us for hundreds of years and because property rights serve an important economic function, facilitating market allocation of scarce goods.

Innovation spillovers are different for three reasons. First, innovations are public goods rather than private goods, which means that they are consumed nonrivalrously and are not naturally scarce: We can all possess and appreciate benefits from an idea without reducing its availability.<sup>56</sup> Ideas can be freely copied by others in the absence of a legal rule

---

absent an exchange. While some uses of land do not deplete it in the same way (though others, such as mining, do), the use of land is rivalrous—if we build a shopping center on a piece of land, you can't use it as a park.

54. See generally Harold Demsetz, *The Private Production of Public Goods*, 13 *J.L. & Econ.* 293, 295 (1970) (finding market price to be most effective mechanism to minimize opportunity cost of private good consumption).

55. To be sure, the scope of a real or personal property entitlement is not always clear. But many of the uncertainties associated with property rights occur precisely when the use of the property is intangible, nonrivalrous, or both—precisely the situations we consider below.

56. Consumption of an idea by one person does not affect the availability of the idea for anyone else; an idea is not depleted in quantity or quality when consumed, regardless of the number of persons consuming it. An idea only needs to be created once to satisfy consumer demand, while an apple must be produced for each consumer. To be clear, some of the resources needed to produce ideas—such as money, time, and human efforts—are scarce, and, as we discuss in the next section, IP rights of one form or another may be necessary in certain circumstances to attract these resources to idea production. Our point here is simply that ideas are naturally “sharable,” which means that efficient allocation of rights to possess and use an idea does not require exclusivity and a market-driven search for the user that values the idea most. Ideas always *can* be shared, possessed,

restricting that copying without depriving their creators of the use of the ideas. Because of that, the transactions that occur in IP are generally not transactions driven by the scarcity of the good and the need to allocate it to the consumer that values it most. Rather, they are driven by the creation of the legal entitlement itself. A patent licensee is not normally buying knowledge she did not otherwise have; she is buying the right not to be sued for using the knowledge she did have, whether because she developed it independently or because she learned it from the inventor's use.<sup>57</sup>

The fact that it is the legal regime itself, not the need for something exclusively in the possession of another, that determines whether a transaction must take place before the buyer can use an idea means that we cannot take the baseline legal regime as a given. If we create strong legal rules governing the use of ideas, doing so will drive a different set of transactions (and innovation-producing activities) than if we make ideas into a commons or create a weaker set of rights in ideas.<sup>58</sup> The externalities that exist in those different sets of transactions cannot be ignored in deciding what legal regime is optimal, even though they are in some sense "internal" to the transaction, because whether there is a transaction at all is determined by society's choice of a legal regime.

The second reason innovation spillovers are different is that many, if not most, of the uses made by "buyers" of IP are productive rather than passive uses—that is, users add to or build on the input in order to create a product or idea that is in turn sold to others. While we can comfortably ignore pecuniary externalities when the buyer is a passive consumer, we cannot do the same with productive uses. Suppose a buyer acquires an input, adds to or improves it or incorporates it into a larger product, and then resells the changed or integrated product. The buyer's valuation of the input is a function of the price at which it can sell the changed product and the cost of the changes it made. But the social value of that input includes not just the price at which the modified good can be sold, but the social benefit others reap from purchasing or using that modified good. If there is consumer surplus in the second transaction—and there

---

and used simultaneously by a wide variety of people in many different ways. Apples generally cannot. For land, it depends upon the context. Whether/when ideas should be shared is a complicated question, which we discuss below, but it is important to recognize the potential in order to appreciate the differences between IP and real and personal property.

57. This is not always true of all forms of IP. Trade secrets are the primary counterexample: If an inventor can keep his idea from becoming public, he can sell access to the idea to a limited number of other people. But even there the trade secret owner generally has to disclose the idea during negotiations so that the buyer knows what he or she is getting. So what the buyer ultimately ends up purchasing is not information he or she didn't know but a legal right to use that information.

58. See Shubha Ghosh, Comment, Competitive Baselines for Intellectual Property Systems, *in* International Public Goods and Transfer of Technology Under a Globalized Intellectual Property Regime 793, 795 (Keith E. Maskus & Jerome H. Reichman eds., 2005).

always is—that consumer surplus is external to the original transaction, because neither the original buyer nor the original seller can capture it. This means that even where the parties can overcome the problems noted above and identify the need for a transaction,<sup>59</sup> we cannot assume that efficient transactions will take place. The buyer’s *private* return to buying, changing, and reselling the good may be less than the seller is charging, even if the *social* return to the changed good would have been enough to justify the transaction.

Third, innovation spillovers are different because the legal rules that define an IP right and determine when transactions must occur lack the clarity of traditional property rights.<sup>60</sup> It is difficult—and in many cases impossible—to know whether one is “trespassing” upon another’s IP right. In part this is a problem with defining the scope of the legal right in question. While courts sometimes talk about patent claims as defining the “metes and bounds” of the right, patent claims lack the certainty associated with real property deeds. Indeed, as Lemley and others have argued elsewhere, the indeterminacy problem is so bad that there may be no way to determine the “right” scope of a patent at all.<sup>61</sup> Not until the Federal Circuit has ruled on the meaning of any particular claim can the patent owner or its competitors know what is owned and what isn’t.<sup>62</sup> And because patent claims are not disclosed to the world for years after

59. In fact, as we discuss below, there are good reasons to be skeptical of the efficient licensing hypothesis. See Frischmann, *Demsetzian Trend*, *supra* note 6 (manuscript at 21–22) (identifying “fundamental flaw” in efficient licensing logic as licensee’s failure to take account of social value of licensee’s use of copyrighted works); Mark A. Lemley, *The Economics of Improvement in Intellectual Property Law*, 75 *Tex. L. Rev.* 989, 1048–72 (1997) [hereinafter Lemley, *Economics of Improvement*] (“Problems of imperfect information, transaction costs, strategic behavior, and market power all impose barriers to the hypothetical efficient license.”); *infra* notes 73–75 and accompanying text.

60. Commenting on earlier drafts of our paper, both Adam Mossoff and Paul Geller suggested that we overstate the clarity of real property. In particular, Mossoff argued that real property boundaries frequently remain uncertain until a conflict between rival users arises, at which point resolution is not always a simple matter. We do not mean (or need) to make a strong claim about the absolute clarity of real property law. Our point is that, comparatively, IP law is far less clear than real and personal property law because of critical differences in the subject matter, the question of whether patents are validly granted, and difficulties that arise in delineating what is being protected.

61. See Dan L. Burk & Mark A. Lemley, *Quantum Patent Mechanics*, 9 *Lewis & Clark L. Rev.* 29, 49 (2005); Craig Allan Nard, *A Theory of Claim Interpretation*, 14 *Harv. J.L. & Tech.* 1, 44 (2000); Margaret Jane Radin, *The Linguistic Turn in Patent Law* 43–48 (2005) (unpublished manuscript, on file with the *Columbia Law Review*), available at <http://justin.hughes.net/ipsc2005/papers/Paper-RADIN.doc>.

62. Even a decision by a district court won’t give significant certainty because the Federal Circuit reverses claim constructions so frequently. The definitive study is Kimberly A. Moore, *Are District Court Judges Equipped to Resolve Patent Cases?*, 15 *Harv. J.L. & Tech.* 1, 2 (2001) (finding that Federal Circuit reverses district court claim constructions thirty-three percent of time). Moore recently updated her study and showed that the problem is getting worse, not better. Kimberly A. Moore, *Markman Eight Years Later: Is Claim Construction More Predictable?*, 9 *Lewis & Clark L. Rev.* 231, 233 (2005); see also Christian A. Chu, *Empirical Analysis of the Federal Circuit’s Claim Construction Trends*,

the invention, there is often no way to have any inkling of whether a technology will end up being patented at all. Trade secrets suffer from a similar indeterminacy problem, though in some sense the problem is even worse, because there is no *ex ante* legal definition of what is and is not a protectable secret. Indeed, virtually no states even require trade secret plaintiffs to define what they claim to own during litigation.<sup>63</sup> Even copyright, which seems the clearest in coverage, can surprise the unwary because of the strict liability requirement.<sup>64</sup>

Nor is the scope of the law any more certain than the scope of particular rights. Users may know that a particular work is copyrighted, but that knowledge gives them little sense of whether a particular use of the work is legal or not, because the idea-expression dichotomy, the filtration of facts and scenes-*a-faire*, the merger doctrine, and the fair use doctrine make it hard to tell whether a surprisingly wide range of uses are permissible. And while patent law does define more or less clearly the legal rights attached to a patent claim, the fact that those rights include the right to prevent all uses of the invention, even those developed independently by the defendant, means that it is effectively impossible for a company making a product to know *ex ante* whether it needs a patent license. Search, identification, and transaction costs are much greater with IP than they are with land or goods. Put another way, IP rights are not—perhaps cannot be—“well-defined” in the sense Coase meant.

The result of these differences between innovation and traditional property regimes is that the universe of pecuniary externalities we can safely ignore is quite a bit narrower in the innovation context than it is with traditional property.<sup>65</sup> While the allocation of benefits between willing participants to most transactions may not matter, the same cannot be said when the government steps in to allocate rights among parties who would not otherwise transact, as it does by creating IP rights. In those cases, the effect a judgment for one side or the other has on the losing party is an important part of the social welfare calculus. So too are the potential effects on third parties who would have benefited from the use forbidden.

In innovation, then, spillovers cannot be ignored. They matter. In the following two sections, we discuss *how* they matter and deconstruct the paradigm of full internalization that the law and economics view offers.

---

16 Berkeley Tech. L.J. 1075, 1096–106 (2001) (finding reversal rate between thirty and thirty-eight percent, depending on assumptions used, and noting that rate was increasing).

63. A rare exception is Cal. Civ. Proc. Code § 2019.210 (West Supp. 2006).

64. See, e.g., *Lipton v. Nature Co.*, 71 F.3d 464, 470–75 (2d Cir. 1995) (holding defendant liable for copyright infringement even though it licensed work from apparent owner because, even though both defendant and owner acted in good faith, the licensor of that owner had apparently obtained it illegally decades before).

65. Cf. Richard A. Posner, *Do We Have Too Many Intellectual Property Rights?*, 9 Marq. Intell. Prop. L. Rev. 173, 174–76 (2005) (describing several distinctions between real and intellectual property).

### B. *We Don't Need Full Internalization of Benefits*

We begin with the supposed need for full internalization of benefits. On the supply side, inventors do not need to capture the full social value of their inventions in order to have sufficient incentive to create.<sup>66</sup> Society needs merely to give them enough incentive to cover the fixed costs of creation that their imitators will not face. Any greater return is at best a mere wealth transfer and at worst wasteful—it doesn't encourage any more innovation in the field, and it may actually interfere with downstream innovation and distort behavior in the market.<sup>67</sup> Thus, while we need some ex ante incentive to innovate, we don't need (and don't particularly want) full internalization of the benefits of an invention. As long as we get enough incentive, the fact that other benefits aren't captured by the innovator doesn't impose any real cost on innovation, and may even contribute to innovation.

Our argument comes in three parts. First, adding reward to a particular innovator does not always encourage more innovation. A firm will make a particular investment in R&D if it expects a positive return on that investment. Unexpected extra return won't affect that investment decision. Expected extra return may or may not encourage still more investment in R&D by that firm.<sup>68</sup> But even if it does, the economic data suggest that invention isn't always a linear, or even monotonic, function of R&D investment. Indeed, smaller companies are often more innovative than larger companies with significantly greater R&D budgets.<sup>69</sup> And

66. See Lemley, *Free Riding*, supra note 20, at 1046–50.

67. *Id.* at 1032 (discussing distortions); Lemley, *What's Different?*, supra note 50, at 1102–03 (explaining that increase in IP rights may encourage new entry, but that entry will be spaced apart by nature of IP rights and therefore will not dissipate returns to IP owners); see also Ghosh, *Rethinking*, supra note 27, at 1357 (“Prospect theorists confuse the inventor’s pursuit of self-interest with the inventor’s appropriation of all the market surplus through strong property rights. The former is what is required for market efficiency, but market efficiency is entirely possible without the latter.” (citation omitted)); Frischmann, *Demsetzian Trend*, supra note 6 (manuscript at 23–24) (discussing externality-induced and property-induced distortions).

68. In an important article, *Externality*, James Buchanan and William Stubblebine differentiate relevant and irrelevant externalities on the basis of the impact that the external costs or benefits have on decisionmaking by relevant agents. James M. Buchanan & Wm. Craig Stubblebine, *Externality*, 29 *Economica* 371 (1962). Simply put, relevant externalities are those costs and benefits for which internalization would lead to a change in behavior (resource allocation); irrelevant externalities are those costs and benefits for which internalization would not lead to a change in behavior (resource allocation). Buchanan and Stubblebine suggest that many externalities are in fact irrelevant and thus fail to justify intervention on efficiency grounds. Those externalities—spillovers—do not need to be internalized. *Id.* at 380–81.

As suggested implicitly by Buchanan and Stubblebine, in many cases, R&D investments will be made with equivalent intensity in settings with and without (some) externalities. In such cases, intervention to increase internalization—through additional property rights or Pigouvian taxes/subsidies—cannot be justified on efficiency grounds.

69. There is a good deal of anecdotal evidence to support the idea that truly innovative ideas come from upstarts rather than established players. See, e.g., Andrew S.

even assuming there is some positive relationship between amount of reward and amount of innovation, that relationship is likely subject to diminishing returns, such that the marginal benefits will at some point be swamped by opportunity costs of investment and the costs of granting more and more control.

Second, those costs of control to innovation are significant. As noted above, full internalization necessarily requires the property owner to control uses and improvements of the innovation. There is significant economic debate over the wisdom of requiring coordinated rather than free market innovation. Kenneth Arrow has argued that it is competition, not central planning (whether by government or a private entity), that most encourages innovation,<sup>70</sup> and while that view is contested, there is empirical evidence to support his argument.<sup>71</sup> If Arrow is correct, full internal-

Grove, *Only the Paranoid Survive* (1996). Much of this evidence is concentrated in the software and Internet industries, where barriers to entry have traditionally been low and where many innovators (Hewlett and Packard, Jobs and Wozniak, Page and Brin) worked outside an established corporate framework.

There is also some limited empirical evidence suggesting that dramatic changes may come from small rather than large companies. See Hearings on H.R. 359, H.R. 632, H.R. 1732, and H.R. 1733 Before the Subcomm. on Courts and Intellectual Property of the H. Comm. on the Judiciary, 104th Cong. 374–83 (1995) (statement of David L. Hill, President, Patent Enforcement Fund, Inc.) (citing Department of Commerce Study from 1960s suggesting that most important inventions were made by independent inventors, even though companies spend more on research and development).

Recent empirical evidence also suggests that newer firms are more innovative than older firms. See Natarajan Balasubramanian & Jeongsik Lee, *Firm Age and Innovation* 14–18 (July 5, 2005) (unpublished manuscript, on file with the *Columbia Law Review*), available at [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=762566](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=762566). This finding seems logically to imply that smaller firms are more innovative than larger ones.

70. Arrow, *Economic Welfare*, supra note 20, at 619. See generally Kenneth J. Arrow, *Collected Papers of Kenneth J. Arrow: The Economics of Information* (1984).

71. See Arrow, *Economic Welfare*, supra note 20, at 620 (“[P]reinvention monopoly power acts as a strong disincentive to further innovation.”); Carliss Y. Baldwin & Kim B. Clark, *Design Rules: The Power of Modularity* (1999) (arguing that fastest way to solve technical problems is to permit independent experiments, not to concentrate resources on one solution); see also Morton I. Kamien & Nancy L. Schwartz, *Market Structure and Innovation* 16 (1982) (discussing various theories of effects of economic structures on rate and form of innovation); F.M. Scherer & David Ross, *Industrial Market Structure and Economic Performance* 660 (3d ed. 1990) (criticizing Schumpeter’s “less cautious” followers for advocating monopoly to promote innovation). In the specific context of IP rights, the canonical argument from both theory and empirical evidence is Robert P. Merges & Richard R. Nelson, *On the Complex Economics of Patent Scope*, 90 *Colum. L. Rev.* 839 (1990); see also Kenneth W. Dam, *The Economic Underpinnings of Patent Law*, 23 *J. Legal Stud.* 247, 252 (1994) (noting that in computer industry, for example, companies coordinate improvements by broad cross licensing because of “the pace of research and development and the market interdependencies between inventions”). For discussions of particular industries in which competition appears to spur innovation, see, e.g., Mark A. Lemley & Lawrence Lessig, *The End of End-to-End: Preserving the Architecture of the Internet in the Broadband Era*, 48 *UCLA L. Rev.* 925, 960–62 (2001) (discussing Internet); Arti Kaur Rai, *Evolving Scientific Norms and Intellectual Property Rights: A Reply to Kieff*, 95 *Nw. U. L. Rev.* 707, 709–10 (2001) (biotechnology); Howard A. Shelanski, *Competition and Deployment of New Technology in U.S. Telecommunications*,

ization not only does not always promote innovation but may actually impede it. Lemley has explained this argument in more detail elsewhere.<sup>72</sup>

Finally, the ability to license others to use or improve on a technology does not necessarily solve this problem. The internalization approach relies on the assumptions that a decisionmaker will rationally act to maximize her own welfare, will be perfectly informed in doing so, and that in doing so, she also will maximize social welfare provided she has internalized the costs and benefits of her action.<sup>73</sup> In reality, none of these assumptions are true. Internalization is costly and far from perfect.<sup>74</sup> While the Coase theorem says that in the absence of transaction costs, rights will be allocated to the highest and best user, in a world with transaction costs we ought to assign property rights to the party best able to minimize costs and maximize benefits, often the best-informed party. Once we admit that we live in a world rife with transaction costs, we must also admit that *both design and allocation of rights matter*.<sup>75</sup> This is especially true of IP.

The assumptions of omniscient rational welfare maximization present even more intractable problems. If a technology might be repurposed or improved in ways that the initial innovator is unlikely to foresee, full internalization will interfere with the socially optimal development and use of a technology. Nor is the problem limited to transaction costs; there are a number of other reasons why efficient licensing may not occur, ranging from strategic behavior to imperfect information<sup>76</sup> to the incentive of an initial innovator to discourage technologies that might compete with it and over which it does not have a comparative advan-

---

2000 U. Chi. Legal F. 85, 85, 96–117 (telecommunications); cf. Frischmann, *Economic Theory*, supra note 3, at 925–26 (observing that most infrastructure resources are “managed in an openly accessible manner to allow” for widespread competitive use rather than coordinated use).

72. See Lemley, *Free Riding*, supra note 20, at 1060–61.

73. As Henry Smith puts it, “The owner then has an incentive to maximize the value of the asset, and, to the extent that internalization has been successful, the owner’s maximization of private value will at the same time maximize the social value of the asset.” Smith, supra note 33, at 1755.

74. As Epstein notes:

[Demsetz’s] account is incomplete for a number of important reasons. In particular, it assumes that the proper regime of property rights can eliminate all externalities. But the real trade-offs are in fact more problematic. The total elimination of these externalities is always an impossibility: externalities occur in too many forms and styles.

Epstein, supra note 1, at S520 (footnote omitted).

75. See Ronald H. Coase, *Prize Lecture to the Memory of Alfred Nobel* (Dec. 9, 1991), available at [http://nobelprize.org/nobel\\_prizes/economics/laureates/1991/coase-lecture.html](http://nobelprize.org/nobel_prizes/economics/laureates/1991/coase-lecture.html) (on file with the *Columbia Law Review*); Lemley, *Free Riding*, supra note 20; Lemley, *What’s Different?*, supra note 50.

76. Lemley, *Economics of Improvement*, supra note 59, at 1048–72.

tage.<sup>77</sup> Moreover, as we discuss in the next section, even absent any of these problems efficient licensing may not occur where licensees' demand for access and use of a technology understates societal demand.

In short, there are strong reasons to be skeptical that an additional increment of reward will necessarily lead to an additional increment of innovation. The relationship is complex, and may even be negative at some point, as the marginal benefits diminish and the central nature of control interferes with innovation by others. We do not need, therefore, to internalize all the benefits of innovation—just enough benefits to encourage the optimal level of innovation.

### C. *We Don't Want Full Internalization of Benefits in All Cases*

On the demand side, arguments for internalization center on the idea that spillovers distort the market allocation of resources because external benefits and costs are not being taken into account by relevant decisionmakers. The theory is that spillovers disrupt the efficacy of the price mechanism and “deprive producers of the signals of consumer preference that trigger and direct their investments.”<sup>78</sup>

There is a fundamental flaw in this logic: The market system works rather well in responding to consumer preferences measured by consumers' willingness to pay for goods and services. But the demand-signaling function of the price mechanism does not necessarily work well when purchasers use a resource as an input to produce a different good for which they themselves cannot expect to capture the full social value. For example, where an input is used to produce public or merit goods, the productive user may fail to observe, appreciate, or appropriate the social value.<sup>79</sup> Purchasers' willingness to pay reflects their private demand—that is, the value that they expect to realize—and does not take into account value that others might realize as a result of their use. To examine this dynamic, Frischmann has focused extensively on “infrastructural” resources—shareable resources capable of being widely used for productive purposes—for which social demand for access and use generally exceeds private demand by a substantial margin.<sup>80</sup> Examples of such goods include education and, significantly for our purposes, information.

77. See, e.g., Arrow, *Economic Welfare*, supra note 20, at 619–21 (noting that monopolists have less incentive to innovate than new entrants because so much of market for new innovation will be cannibalized from monopolist's existing share).

78. Goldstein, supra note 29, at 178–79.

79. See Frischmann, *Economic Theory*, supra note 3, at 962–67; Frischmann, *Demsetzian Trend*, supra note 6 (manuscript at 22).

80. Infrastructural resources represent a subset of resources for which these arguments are especially strong. The limitations on the efficacy of the price mechanism apply more broadly, however, and “can be extended to any situations where productive (re)use of a resource generates relevant spillovers.” Frischmann, *Demsetzian Trend*, supra note 6 (manuscript at 22 & n.58); see also Julie E. Cohen, *Copyright and the Perfect Curve*, 53 *Vand. L. Rev.* 1799, 1812 (2000) [hereinafter Cohen, *Perfect Curve*] (“Creators of these works cannot appropriate all of the value that they create, and so will tend to

Access to infrastructural resources is not necessarily managed well in a private property rights regime.<sup>81</sup> Private property owners are not necessarily optimal suppliers of infrastructure because they have an incentive to investigate and support only those uses that generate observable and appropriable private returns, which may or may not be the uses with the greatest social value. Users are not necessarily optimal purchasers of access, because if they are productive users—as will often be the case with infrastructure—they do not themselves capture the full social value of their use. Their private willingness to pay accordingly understates the social value of their use. Dynamically, this demand manifestation problem works its way upstream and biases allocation, design, investment, and other supply-side decisions concerning the infrastructure resource. As Frischmann puts it:

The market mechanism exhibits a bias for outputs that generate observable and appropriable benefits at the expense of outputs that generate positive externalities. This is not surprising because the whole point of relying on [private] property rights and the market is to enable private appropriation and discourage externalities. The problem with relying on [private property rights and] the market is that potential positive externalities may remain unrealized if they cannot be easily valued and appropriated by those that produce them, even though society as a whole may be better off if those potential externalities were actually produced.<sup>82</sup>

While this bias exists in the presence of monopoly, it also may persist in the absence of supracompetitive pricing. Unlike a reduction in supply associated with the monopoly deadweight loss problem, the problem here is that demand is reduced, in the sense that the demand manifested by productive users falls short of social demand.

Because of these problems, it is not surprising that society regularly manages infrastructural resources—roads, bridges, oceans, the telephone network, and a host of others—in an openly accessible manner to sup-

---

undervalue their uses of preexisting works.”); Lemley, *Economics of Improvement*, supra note 59, at 1056–57 (explaining why reusers cannot pay the full social value of their use); Lydia Pallas Loren, *Redefining the Market Failure Approach to Fair Use in an Era of Copyright Permission Systems*, 5 *J. Intell. Prop. L.* 1, 51–53 (1997) (same); cf. Julie E. Cohen, *Lochner in Cyberspace: The New Economic Orthodoxy of “Rights Management,”* 97 *Mich. L. Rev.* 462, 498 (1998) [hereinafter Cohen, *Lochner in Cyberspace*] (“[T]here is no particular reason to believe that a new author’s ability to pay for the right to use an existing work is a good predictor of the quality of the eventual result, whether quality is measured in terms of market success or by some other standard.”).

81. Frischmann, *Economic Theory*, supra note 3, at 974–78, 986–89; see also Yochai Benkler, *Coase’s Penguin, or, Linux and The Nature of the Firm*, 112 *Yale L.J.* 369, 376–78 (2002) (comparing market-, state-, and commons-based production as information processing systems). See generally Frischmann, *Economic Theory*, supra note 3.

82. Frischmann, *Economic Theory*, supra note 3, at 988–89.

port a wide range of productive activities that generate spillovers.<sup>83</sup> Sometimes, as with the oceans, that management takes the form of a commons—no one owns the resource. Other times, as with most roads and bridges, the government owns the resources and opens access to them. In some cases, such as the telephone network, private ownership of infrastructural resources can help to internalize some externalities (e.g., congestion). Even there, the government intervenes with a common carriage requirement to ensure common access to and use of such resources.<sup>84</sup>

While Frischmann's work focuses on physical infrastructure such as the Internet, roads, and waterways, the core insight about infrastructural management applies to intangible infrastructure, such as ideas, as well.<sup>85</sup> Indeed, intangible infrastructure may be the cleanest example of the benefits of commons because the advantage of private ownership in solving the tragedy of the commons does not apply to information, which is inexhaustible.<sup>86</sup> Ideas themselves are a good example of infrastructure, because they are not merely passively consumed but frequently are reused for productive purposes.<sup>87</sup> There is no reason to believe that ownership of an idea by any one individual will best encourage that productive reuse.<sup>88</sup> Innovation is cumulative and is generally spurred by *decentralized* competition.<sup>89</sup> This is particularly likely to be true of an innovation subject to productive reuse, since no one owner can capture the full value of that innovation anyway. If a company can lock up an invention entirely, the company will have suboptimal incentives to improve it. The owner may (or may not) improve on the invention, but no one else will be in a position to do so.

83. See Carol Rose, *The Comedy of the Commons: Custom, Commerce, and Inherently Public Property*, 53 U. Chi. L. Rev. 711, 719 (1986).

84. Of course, there are other institutional mechanisms besides private property for internalizing externalities associated with congestion (or tragedy of the commons). The difficult policy problem for infrastructure is sustaining a commons while controlling congestion and ensuring supply. There are no easy or universal solutions.

85. See Frischmann, *Commercializing University Research Systems*, supra note 51, at 163–64; Frischmann, *Economic Theory*, supra note 3, at 990–1003; Brett M. Frischmann, *Peer-to-Peer Technology as Infrastructure: An Economic Argument for Retaining Sony's Safe Harbor for Technologies Capable of Substantial Noninfringing Uses*, 52 J. Copyright Soc. 329, 337–43 (2005); Brett M. Frischmann, *The Law and Economics of Ideas, Principles, and Other Intellectual Infrastructures* 6–13 (2005) (unpublished manuscript, on file with the *Columbia Law Review*) [hereinafter Frischmann, *Law and Economics of Ideas*].

86. On the problems with trying to fit the tragedy of the commons into information in order to justify IP protection, see Mark A. Lemley, *Ex Ante Versus Ex Post Justifications for Intellectual Property*, 71 U. Chi. L. Rev. 129, 141–45 (2004) [hereinafter Lemley, *Ex Ante*].

87. See Frischmann, *Economic Theory*, supra note 3, at 990–1003.

88. *Id.*

89. See supra note 71 (discussing literature on competition versus monopoly as spur to innovation).

Frischmann's organizing heuristic is "if infrastructure, then commons."<sup>90</sup> Commons benefit society in circumstances in which they free up the use of ideas that are otherwise subject to underutilization or distorted utilization that does not maximize social value. One cannot automatically make the infrastructure-commons equation, particularly where IP is concerned, since producers need some incentive to innovate. Still, infrastructure theory gives us powerful demand-side reasons for incorporating and sustaining commons within IP rights systems, and therefore for refusing to try to achieve full internalization of spillovers.<sup>91</sup>

#### D. *Combining the Insights*

Combining these supply-side and demand-side ideas gives us an alternative way to think about the relationship between ownership and innovation, one that explains the positive role of spillovers. Many users of innovation are not passive—they incorporate a new innovation into products, build new innovations on top of it, and use it in ways that never occurred to its inventor. Often such productive use yields social value not easily internalized or captured fully by users. A commons, not unitary control, best facilitates this productive reuse. For this reason, IP is a mixed system of private rights and commons—a semicommons—designed to generate both incentives and externalities.<sup>92</sup>

---

90. Frischmann, *Economic Theory*, *supra* note 3, at 922–23 ("My thesis is that if a resource can be classified as infrastructure according to the economic criteria set forth in . . . this Article, then there are strong economic arguments that the resource should be managed in an openly accessible manner.").

91. On infrastructure theory and the economic arguments for managing infrastructure as commons, see generally *id.* at 933–38.

92. As Frischmann has explained:

Copyright is a system designed to both internalize and promote externalities. While this view may seem in direct conflict with the view of copyright as a system designed to create incentives, it is not. Incentives and externalities simply are not mutually exclusive; more of one does not necessarily lead to less of the other, and copyright law promotes both.

Frischmann, *Demsetzian Trend*, *supra* note 6 (manuscript at 5) (citation omitted). Giovanni Ramello explains this point as follows:

[T]he main function of intellectual property is not so much to permit exchanges of intermediate goods (inputs), but rather to cause information to be created and disclosed, thereby facilitating a more complex dynamic connected with the special attributes of knowledge. In a sense, the existence of information externalities is the statutory goal of intellectual property: because information is a public good, the likelihood of opportunistic behaviour within this context is amplified and thus IPRs should address this problem. In the absence of specific property rights, any information will inevitably be followed by its near-immediate imitation, without any remuneration for the creator. The function of IPRs, and the specific economic incentives to creators they provide, is to curtail such occurrences and so avoid underproduction of the market-oriented information, though only through a limited form of appropriability, which should therefore also assure a sufficiently broad spillover effect.

Giovanni B. Ramello, *Property Rights, Firm Boundaries and the Republic of Science: A Note on Ashish Arora and Robert Merges*, 14 *Indus. & Corp. Change* 1195, 1197–98

If innovations were not subject to ownership at all, users would be unconstrained in their ability to make whatever use they wanted of the innovations. In such a state, all the benefits of an innovation would be spillovers. Of course, if innovators could not capture any of the external benefits of their inventions, they would have little or no incentive to invest resources in making those inventions. While some people would invent nonetheless,<sup>93</sup> the inventions such a world would produce would be ones that either took relatively little investment to make or were supported in some other way, such as by a university or a government grant. To encourage private investment into research and development, society must enable investors to capture some of the benefits from that research. This is one role of IP law.

But as Lemley's prior work shows, "some" is not "all."<sup>94</sup> If we calibrate IP rights properly, the disparity between the incentive needed to motivate research and the full social benefit of an innovation will show up as spillovers. And where innovation is subject to productive reuse—and much of it is—Frischmann demonstrates that attempting to capture those spillovers for the innovator through a property right may reduce social welfare in ways that are significant, although difficult to quantify. Spillovers, then, are generally desirable to the extent that we don't need to internalize them in order to get innovation in the first place.

It is of course impossible to know whether society has the "right" level of spillovers. We might have more IP protection than we need to encourage innovation, and too few spillovers. Or it might be that we don't have enough protection, and that expanding IP rights would give society more than enough new innovation to compensate for the inefficiencies caused by restricting use of ideas once they are developed. Or it might be that in some areas we have too much protection, and in other areas we have too little. These are incredibly difficult empirical questions.

---

(2005); see *id.* at 1198 ("The laws for the protection of intellectual property are (or ought to be) designed to properly solve the trade-off between the privatization of information and the preservation of information externalities."); see also Robert A. Kreiss, *Accessibility and Commercialization in Copyright Theory*, 43 *UCLA L. Rev.* 1, 14–20 (1995) (discussing role of copyright law in creating commercial incentives for authors to disseminate new works).

For an analysis of fair use that advocates a broader, more contextual approach to balancing interests within copyright law, see Michael J. Madison, *A Pattern-Oriented Approach to Fair Use*, 45 *Wm. & Mary L. Rev.* 1525, 1622–42 (2004) (arguing that boundaries of copyright law should be defined by "recognized social or cultural" patterns).

93. For example, people invent for reasons such as prestige, because they are paid to do so through government grants or university funds, or because they can find nonlegal ways of appropriating their invention, such as through secrecy or first-mover advantages.

94. See Lemley, *Free Riding*, *supra* note 20, at 1057 ("The broader the scope of an intellectual property right, the less room there is for new innovators to develop and market new products, because the law itself restricts that competition."); Lawrence Lessig, *Intellectual Property and Code*, 11 *St. John's J. Legal Comment.* 635, 638 (1996) ("'Sufficient incentive' . . . is something less than 'perfect control.'").

While we can't identify the optimal level of spillovers, on our approach we can be reasonably confident that optimal level is greater than zero. Spillovers are good for society unless it happens that the incremental social benefits of eking out every last ounce of control outweigh the marginal social cost of that last increment of control. Such a perfect correspondence seems highly unlikely.<sup>95</sup> It turns out, therefore, that internalization of externalities is not the silver bullet that Demsetzians imagine. And as noted above, there is good economic evidence that greater innovation spillovers are associated with more, not less, innovation,<sup>96</sup> which suggests that if anything we have too much IP protection and too few spillovers today. Traditional law and economics approaches to IP cannot explain why there should ever be such a relationship between spillovers and productivity. Our new approach does explain it, and lays the groundwork for a new way of thinking about the economics of IP.

### III. APPLICATIONS TO IP AND THE INTERNET

Having laid the foundation for an alternative economic theory of IP and spillovers, in this Part we explore a few of the many possible applications of the theory to copyright and patent law and to the debate over network neutrality.<sup>97</sup>

#### A. Copyright

Copyright is a system designed to generate both incentives and spillovers. Copyright law creates a semicommons arrangement—a complex mix of private rights and commons.<sup>98</sup> The rights granted by copyright law—specifically, the § 106 rights to reproduce, display, perform, distribute, and make derivative works<sup>99</sup>—provide incentives to create and disseminate works by facilitating transactions and lowering the costs of excluding competitors from using the expression. The supply-side incentives that copyright affects extend beyond the initial investment in creation to investments in development and dissemination of content. It is not just creation but publication, dissemination, and productive use of works that must also be encouraged. Like traditional property rights, copyright facilitates transactions over certain uses of creative expression.

---

95. Indeed, even some advocates of increased control base their argument on the continued presence of spillovers, assuming that increased control will still leave us with a “leaky” intellectual property system. See generally R. Polk Wagner, *Information Wants to Be Free: Intellectual Property and the Mythologies of Control*, 103 *Colum. L. Rev.* 995 (2003) (arguing that “control” conferred by intellectual property rights does not actually reduce availability of information in public domain).

96. See *supra* notes 5–8, 34–49 and accompanying text.

97. Other applications can readily be imagined, ranging from the desirability of the Digital Millennium Copyright Act to the role of subject matter limitations in patent law. Our goal in this section is merely to identify a few significant applications of our theory.

98. Portions of this and the next paragraph are adapted from Frischmann, *Demsetzian Trend*, *supra* note 6.

99. 17 U.S.C. § 106 (2000).

It thereby enables rightsholders to appropriate some of the surplus generated by their investments in creation, development, and dissemination. In this fashion, the private rights component of copyright law improves investment incentives through the operation of the market mechanism, in a sense using the market to achieve a broader set of economic and social ends.

The commons component of copyright law promotes spillovers. Through a variety of leaks and limitations on the private rights granted, copyright law sustains common access to and use of resources needed to participate in a wide variety of intellectually productive activities.<sup>100</sup> Many of these activities generate socially valuable spillovers: value realized by consumers, users, and third parties that is external to a creator's decision to produce the work and to any transactions involving the work. For example, due to its limited duration,<sup>101</sup> copyright has generated temporal externalities. A work that enters the public domain is free for public use, and the value derived from such use is external to the creator's decision to produce the work and to any transactions involving the work. Similarly, due to copyright's limited scope,<sup>102</sup> copyright generates externalities that accrue to other creators, even competitors, as these entities can freely use various unprotected elements of a work, such as an idea, theme, or functional feature. Copyright's limited scope may also generate externalities in complementary technology markets as companies design and build products like DVD players and iPods that facilitate the enjoyment of copyrighted works.<sup>103</sup> Finally, copyright produces externalities when consumers make productive (re)use of the works they consume. Creating and consuming creative expression of different types develops human capital, educates, and socializes in a manner that benefits not only creators and consumers but also nonparticipants.<sup>104</sup> We discuss a few examples.

---

100. See, e.g., Frischmann, *Economic Theory*, supra note 3, at 1000–03 (“Intellectual property is designed, at least in theory, to promote and preserve a sustainable information commons.”); Kreiss, supra note 92, at 4 (“To function properly, copyright law must strike a balance between the rights given to copyright authors and the access given to copyright users.”); cf. Wagner, supra note 95, at 997 (asserting that control critics overstate degree to which creators' control of intellectual property rights decreases information in public domain).

101. See U.S. Const. art. I, § 8, cl. 8 (“Limited Times”); 17 U.S.C. § 302 (establishing duration of copyright protection).

102. See, e.g., 17 U.S.C. §§ 102, 107 (providing for limits on scope of copyrightable subject matter and fair use defense). See generally William M. Landes & Richard A. Posner, *The Economic Structure of Intellectual Property Law* (2003) (discussing the purposes of limits on copyright).

103. Fred von Lohmann, *iPods, TiVo and Fair Use as Innovation Policy* (Mar. 31, 2005) (unpublished manuscript, on file with the *Columbia Law Review*).

104. See Cohen, *Perfect Curve*, supra note 80, at 1812; Frischmann, *Economic Theory*, supra note 3, at 1003; Loren, supra note 80, at 49; cf. Frischmann, *Commercializing University Research Systems*, supra note 51, at 169–71.

1. *Subject Matter Exclusion: Ideas.* — Copyright law extends protection to expression of an idea but categorically excludes from protection the idea itself.<sup>105</sup> Ideas are intellectual infrastructure: They are consumed nonrivalrously and can be used productively to produce a wide variety of outputs. Ideas generally may be expressed and communicated in many different ways, and it is only the particular expression that copyright law protects.<sup>106</sup> Copyright does not extend to ideas because propertizing ideas, first, is not generally necessary to create incentives to invest resources in idea production and, second, would inhibit productive use of ideas and the spillovers that would flow from such use.

In addition to excluding some subject matter outright, copyright law also excludes elements of a protected work and sustains common access to those elements. For example, the historical facts expressed in a book can be copied and reused without a license.<sup>107</sup> Generally, facts are said to be excluded from copyright law because they do not originate with the author but are mere discoveries.<sup>108</sup> Putting aside attempts to distinguish creation from discovery,<sup>109</sup> an important economic reason for excluding facts from copyright protection is that facts are essential inputs into socially valuable processes of learning about, experiencing, and understanding our world. Facts about the world are generally excluded because of their essential role in socially valuable processes, like education, that generate spillovers.

2. *Defenses: Fair Use.* — Fair use is a defense to copyright infringement that preserves a range of unlicensed uses of copyright protected expression. It is critical to the semicommons design and utilitarian objectives of copyright law.<sup>110</sup> The defense originated and developed through

105. Copyright protection does not extend to “any idea, procedure, process, system, method of operation, concept, principle, or discovery.” 17 U.S.C. § 102(b); see also *Baker v. Selden*, 101 U.S. 99, 101–03 (1879) (holding that copyright for book describing bookkeeping system does not give copyright holder exclusive right to use bookkeeping system); *Nichols v. Universal Pictures Corp.*, 45 F.2d 119, 121–22 (2d Cir. 1930) (finding that later play did not infringe copyright of earlier play because copyright holder could not copyright general theme of play).

106. Distinguishing between idea and expression is not as simple as it sounds and has proven to involve difficult line drawing. See *Baker*, 101 U.S. at 101–03.

107. See *Hoehling v. Universal City Studios, Inc.*, 618 F.2d 972, 979–80 (2d Cir. 1980) (holding that historical facts are in public domain and that “[i]n works devoted to historical subjects . . . a second author may make significant use of prior work, so long as he does not bodily appropriate the expression of another”), cert. denied, 449 U.S. 1 (1980).

108. *Feist Publ’ns Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 347 (1991).

109. Both discovery and creation involve work, an investment of resources, and thus both may require inducement to invest to overcome risks of free riding.

110. As Pierre Leval explains:

The doctrine of fair use limits the scope of the copyright monopoly in furtherance of its utilitarian objective. . . . Fair use should not be considered a bizarre, occasionally tolerated departure from the grand conception of the copyright monopoly. To the contrary, it is a necessary part of the overall design. Although no simple definition of fair use can be fashioned, and inevitably

common law decisions and was codified in 1976 as an open-ended, multifactor test.<sup>111</sup> The purposes listed are illustrative, and the four factors are not exclusive. Application of fair use is thus a fact-intensive inquiry engaged in on a case-by-case basis.

The traditional law and economics explanation of fair use focuses on transaction costs.<sup>112</sup> The explanation is rather simple—when transaction costs are prohibitive, an efficient deal will not be struck and uses that should happen do not happen. So fair use sidesteps dealmaking altogether and leaves the public free to engage in the uses without transacting with the copyright owner. Our theory makes clear that the transac-

---

disagreement will arise over individual applications, recognition of the function of fair use as integral to copyright's objectives leads to a coherent and useful set of principles. Briefly stated, the use must be of a character that serves the copyright objective of stimulating productive thought and public instruction without excessively diminishing the incentives for creativity.

Pierre N. Leval, *Toward a Fair Use Standard*, 103 Harv. L. Rev. 1105, 1110 (1990).

111. The test was codified as follows:

Notwithstanding the provisions of sections 106 and 106A, the fair use of a copyrighted work, including such use by reproduction in copies or phonorecords or by any other means specified by that section, for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright. In determining whether the use made of a work in any particular case is a fair use the factors to be considered shall include—

- (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
- (2) the nature of the copyrighted work;
- (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
- (4) the effect of the use upon the potential market for or value of the copyrighted work.

The fact that a work is unpublished shall not itself bar a finding of fair use if such finding is made upon consideration of all the above factors.

17 U.S.C. § 107 (2000).

112. According to Matthew Sag, the “two primary law and economics contributions to the fair use doctrine [are] the market failure test and the application of a cost-benefit analysis.” Sag, *Beyond Abstraction*, supra note 31 (manuscript at 30). The market failure approach to fair use embodies the transaction cost story we discuss in the text. See, e.g., Landes & Posner, supra note 102, at 358–59; Wendy J. Gordon, *Fair Use as Market Failure: A Structural and Economic Analysis of the Betamax Case and Its Predecessors*, 82 Colum. L. Rev. 1600 *passim* (1982); William F. Patry & Richard A. Posner, *Fair Use and Statutory Reform in the Wake of Eldred*, 92 Cal. L. Rev. 1639, 1646 (2004). The cost-benefit approach, which is more akin to our own, would determine whether a use should be deemed fair based on an analysis of the social costs and social benefits. See, e.g., William W. Fisher III, *Reconstructing the Fair Use Doctrine*, 101 Harv. L. Rev. 1661, 1705–17 (1988) (outlining considerations judges should take into account in fair use conflicts based on comprehensive cost-benefit analysis); Glynn S. Lunney, *Fair Use and Market Failure: Sony Revisited*, 82 B.U. L. Rev. 975, 996 (2002) (“[F]air use must . . . entail a balancing of the public benefits and losses associated with granting the copyright owner the right to prohibit particular uses.”).

tion cost explanation only captures a portion of the fair use space and does not, and should not, fully define fair use.<sup>113</sup>

The transactions we tend to think about when discussing fair use and the transaction cost story tend to be transactions between the copyright owner and the user, and in evaluating the story (are transaction costs high or low?), we tend to retain our focus on these parties. This narrow focus misses spillover effects and fails to consider the possibility that the user may decide not to seek a license or may take a license under conditions or with restrictions that reduce or eliminate benefits that spill over to third parties. Fair use deems lawful some uses that yield benefits to third parties, not because the transaction costs between the copyright owner and user are necessarily high, but rather to sustain the flow of spillovers to third parties. As we explained above, in some contexts the user's willingness to pay for access to and use of the copyrighted material may understate social demand for that party's access and use because the licensee fails to observe, appreciate, or capture spillovers that flow to third parties.<sup>114</sup>

The spillover benefits realized individually may be small, which perhaps explains why the spillover benefits tend to be ignored in a transaction cost analysis.<sup>115</sup> But in some cases, the spillover benefits may add up to a considerable amount of value, for example, when small-scale benefits are widely dispersed within a community.<sup>116</sup> Many paradigmatic uses deemed fair involve use of a work to engage in activities that yield diffuse,

113. We are certainly not the first to make this point. See, e.g., Lunney, *supra* note 112, at 993. Indeed, Wendy Gordon, who pioneered the transaction cost theory of fair use, has recently emphasized that it explains only one subset of fair use cases. See Wendy J. Gordon, *Market Failure and Intellectual Property: A Response to Professor Lunney*, 82 *B.U. L. Rev.* 1031, 1034 (2002).

114. See, e.g., Dan L. Burk & Julie E. Cohen, *Fair Use Infrastructure for Rights Management Systems*, 15 *Harv. J.L. & Tech.* 41, 44–45 (2001) (discussing 2 Live Crew parody case as example where “value of socially beneficial uses of copyrighted works is not fully internalized” (citing *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569 (1994))); Cohen, *Lochner in Cyberspace*, *supra* note 80, at 545 (arguing that information has “ancillary social value” and that markets generally do not recognize that value); Lemley, *Economics of Improvement*, *supra* note 59, at 1056–58 (discussing possibility of underproduction of goods with positive externalities).

115. While transaction costs could be defined to encompass any barrier that would preclude reaching a socially optimal outcome, such a broad vision of transaction costs conflates much too much. We could say that when a user fails to capture spillovers that result from his use of copyrighted expression, transaction costs in a secondary “market” are prohibitive—the user cannot capture the benefits in a transaction with the people who realize the spillover benefit. The problem with this view is that in many, if not most, cases no transaction is even possible. This might seem like an argument that wins the day for fair use because transaction costs are so high and cannot be reduced. But that would go too far—the mere existence of a use that generates third-party benefits over which no transaction will take place does not mean that the use should be deemed fair. The transaction cost trigger is not determinative in this context.

116. See Frischmann, *Economic Theory*, *supra* note 3, at 975–76 (explaining how widespread production of small social surpluses may add up to significant social windfall).

small-scale spillovers to a community. Using a work for educational purposes, for example, not only benefits the users themselves, but also, in a small way, benefits others in the users' community with whom users have interdependent relations—reading and learning builds socially valuable human capital. Critiquing a work similarly benefits not only the user but also, in a small way, others in the users' community—not only because those others may read the critique itself, but also because engaging in critical commentary is a form of creative and cultural activity that builds socially valuable human capital.

We recognize that observing and measuring these spillover benefits is probably an impossible task. That is our point, in fact. As a society, on the whole, we recognize the value of active, widespread participation in these types of activities, and we know that creative expression is essential to participation. Thus, we encourage common access to and use of expression for these types of activities. Doing so provides a justification for a fair use doctrine based on public benefit (part of the first factor of the fair use analysis) and not just revenue maximization to the copyright owner.<sup>117</sup>

Of course, in determining whether a particular use should be deemed fair under copyright law, the benefits of fair use are only part of the equation; the costs matter as well. The potential costs of fair use are the reduction in incentives to invest in the creation, development, management, or dissemination of creative expression.<sup>118</sup> Various factors in the fair use analysis act as proxies for incentive-based costs. For example, the first fair use factor takes into account whether a use is commercial or noncommercial.<sup>119</sup> Deeming commercial uses fair obviously would have a comparatively greater impact on incentives than deeming noncommercial uses fair. The commercial nature of a use also suggests greater potential for transactions to internalize some or all of the benefits in question. Even more on point, the fourth factor takes into account the effect that a use has on existing or potential markets for the copyrighted work.<sup>120</sup> Here the analysis is directly aimed at assessing the impact on the copyright owner of deeming a use fair. Courts ask whether the defendant's use leads to a substitute expression that will compete directly with the original work being used without permission or, alternatively, with deriva-

---

117. See Leval, *supra* note 110, at 1110 (espousing fair use standard “that serves the copyright objectives of stimulating productive thought and public instruction”).

118. While it might be tempting to load the costs of administering fair use on to the pile, there is no reason to allocate such costs exclusively to the “costs of fair use” rather than the “costs of unfair use.” The line drawing accomplished by the fair use doctrine is merely one of many examples of how copyright law allocates private and common rights. In a systematic assessment of how copyright law allocates these rights, the administrative costs surely matter and must be accounted for, but in an analysis of whether a particular use of a work should be deemed fair, administrative costs are not relevant.

119. 17 U.S.C. § 107(1) (2000).

120. *Id.* § 107(4).

tives of the original.<sup>121</sup> To the extent that substitution is likely, there is likely a greater impact on incentives, and this is a social cost to deeming the use fair. If market substitution is unlikely, however, the risk to incentives is smaller.

This calculus is threatened by the notion that an independent market for licenses exists, such that a use that might be feasibly licensed ought to be deemed unfair because deeming such a use fair would deprive the copyright owner of license fees and thus diminish incentives.<sup>122</sup> As many have observed, it is circular to rely on the existence of a market for licenses when determining whether a use should be deemed fair, because such a market exists only if the use is deemed unfair (or where uncertainty about the status of the use makes licensing attractive).<sup>123</sup> The market for licenses is dependent upon the lines drawn by the law.<sup>124</sup> This circular reasoning is a problem because it fully defines the fair use space according to transaction costs and ignores the benefits of spillovers. Some uses should be permitted because of the social benefits they create, or because copyright owners don't need the unanticipated extra revenue, even if permitting the uses "costs" the copyright owner revenues he or she might have been able to collect if the law were different.

## B. Patent

Patent law, like copyright, is a semicommons that promotes both ownership of rights and spillovers, though the particular ways in which patent law permits "leakage" differ significantly from copyright law. Patent law protections have a much shorter duration than copyright, permitting inventions to enter the public domain more quickly. Patent law also excludes some inventions from protection because requirements for obtaining protection are stricter. Once inventors do obtain protection, however, they obtain a right that is much stronger and less leaky than that afforded by copyright law.

---

121. See, e.g., *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 590–94 (1994) (analyzing market harm that rap parody allegedly inflicted on "Oh, Pretty Woman" and its possible derivatives).

122. For a discussion of these cases, and criticism of them along different lines, see Mark A. Lemley, *Should a Licensing Market Require Licensing?*, *Law & Contemp. Probs.* (forthcoming 2006) (manuscript at 4–11, on file with the *Columbia Law Review*).

123. See, e.g., Christina Bohannon, *Reclaiming Copyright*, 23 *Cardozo Arts & Ent. L.J.* 567, 597–98 (2006) (observing "confusion, even circularity, in the fair use analysis"); Fisher, *supra* note 112, at 1671 (revealing circularity in Supreme Court's definition of fair use); Leval, *supra* note 110, at 1124 ("By definition every fair use involves some loss of royalty revenue because the secondary user has not paid royalties." (footnote omitted)); Loren, *supra* note 80, at 38–39 (noting "[i]nherent [c]ircularity" of courts' consideration of "lost" permission fees in analyzing fairness of use); Matthew Sag, *God in the Machine: A New Structural Analysis of Copyright's Fair Use Doctrine*, 11 *Mich. Telecomm. & Tech. L. Rev.* 381, 393–95 (2005) (discussing economics of fair use); Sara K. Stadler, *Forging a Truly Utilitarian Copyright*, 91 *Iowa L. Rev.* 609 (2006) (noting problem with baselines in fair use cases).

124. See *supra* notes 56–58 and accompanying text.

Patent law promotes spillovers in several ways. Patents generate externalities by facilitating learning and disclosure. Indeed, patent law (unlike copyright law) requires the patent owner to teach the public how to make and use the invention, and this is often identified as a central function of the patent system, though in practice it is considerably less important than the incentive effects of the system. Patents lead to temporal externalities—spillovers upon the expiration of the patent. Because inventions tend to have more staying power than creative works, those temporal spillovers are quite significant. The overwhelming majority of the social benefit associated with the telephone (and, for that matter, the paper clip) occurred after the basic patents on those technologies expired.

Recognizing that patent law, like copyright law, is designed not simply to maximize inventor return but also to promote spillovers has implications for several important issues currently in dispute in patent law.

1. *Scope of Rights: Doctrine of Equivalents.* — Patents literally cover the scope defined in their claims. But courts long ago recognized that limiting patent scope to the literal terms of the claims could sometimes lead to unfair results, permitting a defendant to avoid liability by making an insignificant change to the invention.<sup>125</sup> Patent law's doctrine of equivalents aims to avoid this problem by allowing patent owners to capture within the scope of their rights products that are only insubstantially different from the invention as described.<sup>126</sup> A major application of the doctrine of equivalents is to later-developed technologies—things the patent owner could not have foreseen at the time the patent was written.<sup>127</sup> The question becomes how much control patent owners should have over not merely foreseeable but unforeseeable developments or improvements of their invention.

---

125. See, e.g., *Graver Tank & Mfg. v. Linde Air Prods. Co.*, 339 U.S. 605, 607–09 (1950) (“[T]o permit imitation of a patented invention which does not copy every literal detail would be to convert the protection of the patent grant into a hollow and useless thing.”).

126. *Id.*; see also *Warner-Jenkinson Co. v. Hilton Davis Chem. Co.*, 520 U.S. 17, 40–41 (1997) (refusing to discard doctrine of equivalents).

127. For a discussion of this problem, see, e.g., *Sage Prods., Inc. v. Devon Indus.*, 126 F.3d 1420, 1424–25 (Fed. Cir. 1997) (refusing to apply doctrine of equivalents expansively because “[s]ociety at large would bear these . . . costs in the form of virtual foreclosure of competitive activity within the penumbra of each issued patent claim”); Matthew J. Conigliaro et al., *Foreseeability in Patent Law*, 16 *Berkeley Tech. L.J.* 1045, 1047 (2001) (weighing costs and benefits of doctrine of equivalents); Doug Lichtman, *Substitutes for the Doctrine of Equivalents: A Response to Meurer and Nard*, 93 *Geo. L.J.* 2013, 2025–27, 2030–32 (2005) (defending doctrine of equivalents for deterring malevolent copyists, lowering costs of patent claim writing, and permitting more informed determinations about patent scope). But see Michael J. Meurer & Craig Allen Nard, *Invention, Refinement, and Patent Claim Scope: A New Perspective on the Doctrine of Equivalents*, 93 *Geo. L.J.* 1947, 1996–98 (2005) (arguing against application of doctrine of equivalents to later-developed technology because doctrine will not lower costs of patent applications or increase incentives to invent).

Thinking about later-developed technologies in terms of spillovers may help us to understand that question. As with copyright and fair use, the question is one of social costs and benefits. If an invention is too easy to evade, patent rights will not create strong incentives to invent, particularly for the most important inventions.<sup>128</sup> On the other hand, our discussion of the value of spillovers suggests that patent owners should not have a claim to capture all the social value attendant upon their invention. With changed circumstances and over time, more and more of that value will accrue to outsiders who tinker with or repurpose the invention. That's a desirable feature of the system, not a bug. The doctrine of equivalents, then, must not be seen as a license to sweep any application that derives value from the patented invention into the control of the patent owner. Rather, the cost-benefit analysis of the doctrine of equivalents as applied to later-developed technologies should balance the incremental incentive created by the prospect of a broad right extending into unforeseen territory against the potential that broad application of the doctrine will restrict improvers who cannot effectively bargain for those rights.<sup>129</sup>

2. *Remedies.* — The proper remedies upon a finding of patent infringement are a matter of great dispute. The Supreme Court's most recent patent decision rejected a longstanding rule that patent owners were always entitled to injunctive relief against infringement, instead applying the traditional four-factor equitable test for deciding whether a court should issue an injunction.<sup>130</sup> And Congress is currently debating patent reform that will address, among other things, the problem of overcompensating patent owners whose rights extend to only one small piece of a larger integrated product.<sup>131</sup> Spillovers can inform this debate as well, in similar ways. Patent owners should not always be entitled to capture the full social benefit of their invention. Rather, particularly in circumstances where the defendant or third parties made significant contributions to the success of the product, social welfare requires that they be entitled to continue to make use of that product. Patent owners should

---

128. On the pioneering invention doctrine, which gives greater scope to more important inventions, but which has fallen into disfavor of late, see *Miller v. Eagle Mfg. Co.*, 151 U.S. 186, 207 (1894) ("If the invention is broad or primary in its character, the range of equivalents will be correspondingly broad, under the liberal construction which the courts give to such inventions."); *Perkin-Elmer Corp. v. Westinghouse Elec. Corp.*, 822 F.2d 1528, 1532 (Fed. Cir. 1987) ("A pioneer invention is entitled to a broad range of equivalents." (citation omitted)); John R. Thomas, *The Question Concerning Patent Law and Pioneer Inventions*, 10 *High Tech. L.J.* 35, 37 (1995) ("Courts construe pioneer patent claims . . . to encompass a broader range of so-called 'equivalents' during an infringement determination." (footnote omitted)).

129. For an effort at this analysis, see Lemley, *Economics of Improvement*, *supra* note 59, at 1048–72.

130. *eBay, Inc. v. MercExchange LLC*, 126 S. Ct. 1837, 1838–39 (2006).

131. See H.R. 2795, 109th Cong. (2005). On this "royalty stacking" problem, see, e.g., Mark A. Lemley & Carl Shapiro, *Patent Holdup and Royalty Stacking*, 85 *Tex. L. Rev.* (forthcoming 2007) (manuscript at 1, on file with the *Columbia Law Review*).

be compensated for the use, to be sure, but in those cases compensation—and not a property right of complete control—is all they should receive. The theory we have developed can also help people to understand why patents are in this respect not like real property, and therefore serve to beat back periodic assertions that denying an injunction to a patent owner is like forcing me to let others use my house when I'm not home.<sup>132</sup> Given the infrastructure at issue with IP, and the differences in boundaries and transactions costs, a better “real-world” analogy might be forcing me to let others use the airspace above my house—something that was once contested<sup>133</sup> but that we now permit as a matter of course.

### C. *Spillovers and Network Neutrality*

To illustrate how our theory might apply outside of IP, we briefly explore in this section the “network neutrality” debate taking place in the field of communications law.<sup>134</sup> At the heart of this debate is whether (1) the Internet will retain its end-to-end architecture and continue to be managed in an openly accessible, nondiscriminatory manner or (2) the owners of the networks that jointly comprise the physical infrastructure of the Internet will be able to discriminate among Internet users and uses in an effort to extract as much of the value realized by users as possible.

The Internet consists of many infrastructure resources.<sup>135</sup> Scholars have delineated two macrolevel infrastructure resources. The *physical in-*

132. See, e.g., F. Scott Kieff & R. Polk Wagner, Op-Ed., Testing Patent Protections, Wash. Times, May 30, 2006, at A14 (analogizing courts' refusal to automatically enjoin patent infringements to “a rule that allowed me, anytime I notice you are not using your car as I like, to use it myself and pay whatever a court might later request, if you sue me and win”).

133. See Lawrence Lessig, Free Culture: How Big Media Uses Technology and the Law to Lock Down Culture and Control Creativity 1–3 (2004).

134. See, e.g., Tim Wu, The Broadband Debate, a User's Guide, 3 J. on Telecomm. & High Tech. L. 69, 88–94 (2004); Michael Powell, Chairman, Fed. Commc'ns Comm'n, Preserving Internet Freedom: Guiding Principles for the Industry, Remarks at the Silicon Flatirons Symposium: Digital Broadband Migration: Toward a Regulatory Regime for the Internet Age 4 (Feb. 8, 2004), available at [http://hraunfoss.fcc.gov/edocs\\_public/attachmatch/DOC-243556A1.pdf](http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-243556A1.pdf) (on file with the *Columbia Law Review*); Letter from Tim Wu, Assoc. Professor, Univ. of Va. Law Sch., & Lawrence Lessig, Professor of Law, Stanford Law Sch., to Marlene H. Dortch, Sec'y, Fed. Commc'ns Comm'n 3 n.3 (Aug. 22, 2003), available at [http://gullfoss2.fcc.gov/prod/ecfs/retrieve.cgi?native\\_or\\_pdf=pdf&id\\_document=6514683885](http://gullfoss2.fcc.gov/prod/ecfs/retrieve.cgi?native_or_pdf=pdf&id_document=6514683885) (on file with the *Columbia Law Review*).

135. Portions of this subsection are adapted from Frischmann, Economic Theory, supra note 3, at 1005–08, and are developed further in Brett Frischmann & Barbara van Schewick, Yoo's Frame and What It Ignores: Network Neutrality and the Economics of an Information Superhighway, 47 Jurimetrics J. (forthcoming 2007) (manuscript at 7–12, on file with the *Columbia Law Review*). While there are important examples of infrastructure at the applications layer of the Internet, such as web browsers, search engines, IM systems, video players, and so on, our focus here is on the underlying physical and logical infrastructures of the Internet. For a discussion focusing on the applications layer and IP law, see generally Philip J. Weiser, The Internet, Innovation, and Intellectual Property Policy, 103 Colum. L. Rev. 534 (2003).

*frastructure* consists of a wide variety of physical networks interconnected with each other, while the *logical infrastructure* consists of the standards and protocols that facilitate seamless transmission of data across different types of physical networks.<sup>136</sup> The physical and logical infrastructure both act as essential inputs into downstream production of applications and content, and thus constitute the foundational layers upon which the Internet environment we experience has been built.<sup>137</sup>

The Internet currently is managed in an openly accessible, nondiscriminatory manner. Network owners interconnect and do not discriminate among each others' customers; nor do they discriminate on the basis of applications or content. The current Internet infrastructure evolved with the so-called "end-to-end" design principle as its central tenet.<sup>138</sup> This design principle is implemented in the logical infrastructure of the Internet through the adoption of standardized communication protocols (e.g., the Internet Protocol suite).<sup>139</sup> In order to preserve the simplicity and robustness of the network and allow applications to be easily layered on top of it, end-to-end architecture prevents infrastructure providers from discriminating among data packets carried by their networks.<sup>140</sup> This design promotes the open interconnection of networks and focuses application development and innovation on the demands of end users.<sup>141</sup> For the most part, infrastructure providers are ignorant of the identity of end users and end uses, and for their part end users are ignorant of the various networks that transport data packets (with the exception of the "last mile" provider with whom they connect). In a sense, shared ignorance is "built" into the infrastructure and precludes individualized exclusion of end users or end uses.

End-to-end design sustains an infrastructure commons by insulating end users from market-driven restrictions on access and use.<sup>142</sup> Because infrastructure providers cannot distinguish between end uses or end users, they cannot base access decisions or pricing on who is sending packets or how those packets may be used; nor can they optimize the infrastructure for a particular class of end uses or end users. Functionally, the end-to-end principle acts as a limitation on the property rights of

---

136. Yochai Benkler, *From Consumers to Users: Shifting the Deeper Structures of Regulation Toward Sustainable Commons and User Access*, 52 *Fed. Comm. L.J.* 561, 562 (2000).

137. See Philip J. Weiser, *Law and Information Platforms*, 1 *J. on Telecomm. & High Tech. L.* 1, 4–5 & n.14 (2002) (describing layered model of Internet).

138. Lawrence Lessig, *The Future of Ideas: The Fate of the Commons in a Connected World* 34–35 (2001) [hereinafter Lessig, *Future of Ideas*]; Lemley & Lessig, *supra* note 71, at 930–31.

139. See Joseph Farrell & Philip J. Weiser, *Modularity, Vertical Integration, and Open Access Policies: Towards a Convergence of Antitrust and Regulation in the Internet Age*, 17 *Harv. J.L. & Tech.* 85, 91 (2003) (describing how Internet Protocol implements end-to-end architecture).

140. Lemley & Lessig, *supra* note 71, at 931–32.

141. See *id.*

142. See Lessig, *Future of Ideas*, *supra* note 138, at 46–47.

network owners, much like fair use operates as a limitation on the rights of copyright owners.

There is considerable pressure for change: pressure to replace the existing “dumb,” open architecture with an “intelligent,” restrictive architecture capable of differentiating and discriminating among end uses and end users; pressure for property rights evolution so that property owners may more fully internalize externalities and appropriate the value of the Internet. This pressure comes from many sources, including the Internet’s evolution to broadband (infrastructure, applications, and content), the demand for latency-sensitive applications such as video-on-demand and IP telephony, the demand for security measures and spam regulation measures that are implemented at the “core” of the Internet,<sup>143</sup> and, more generally and importantly, demand for increased returns on infrastructure investments.

These pressures for change have given rise to the “network neutrality” debate. Owners of physical communications networks have argued for the abandonment of the end-to-end principle and for the freedom to discriminate among users and uses with respect to price or quality of service. Their arguments are essentially Demsetzian. Network owners claim that they should be free to use their property as they see fit and that only by having complete dominion over their networks will they have properly aligned incentives to deploy broadband access. For example, they argue that congestion externalities should be alleviated through a pricing mechanism sensitive to the user or use in question.<sup>144</sup> Further, they argue that there are a host of uncaptured spillover benefits realized by applications providers, content providers, and end users; according to some network executives, these “free-loaders” should be made to pay.<sup>145</sup> It is

---

143. These pressures reflect important concerns and give rise to difficult tradeoffs. Quality of service and security, in particular, are stifled to some degree by a “dumb” architecture. It is worth noting, however, that even with a dumb architecture, innovators have figured out how to provide certain degrees of quality of service at the periphery of the network and how to make certain latency-sensitive applications, such as IP telephony, work.

144. It is worth noting that there are alternative means for dealing with congestion without content discrimination, such as charging end users according to the bandwidth consumption and/or time of day. Congestion-based (or usage-based) pricing need not be discriminatory. Christopher Yoo argues that network owners should be able to use discriminatory pricing systems that rely on applications as proxies for metering usage, Christopher S. Yoo, *Network Neutrality and the Economics of Congestion*, 94 *Geo. L.J.* 1847, 1907–08 (2006), but he fails to make the case for using such proxies and simply conjectures that the costs of metering bandwidth consumption may exceed the costs of metering by application type. Moreover, Yoo myopically focuses on the private costs of metering and fails to consider the social costs of discriminatory pricing systems, which we discuss below. And he is far too sanguine about the disciplinary effects of what limited competition exists in the wireline space. See Frischmann & van Schewick, *supra* note 135 (manuscript at 24–28).

145. According to John Thorne, a Verizon senior vice president and deputy general counsel, “The network builders are spending a fortune constructing and maintaining the networks that Google intends to ride on with nothing but cheap servers. . . . [Google] is enjoying a free lunch that should, by any rational account, be the lunch of the facilities

not that these parties are getting free access to the network; indeed, all three groups pay handsomely for broadband Internet access, generating a major source of revenue for network owners. Rather, network owners complain that the inability to discriminate among users and uses precludes them from extracting a still greater share of the surplus.<sup>146</sup> Much like the arguments raised in favor of strengthening IP rights, network owners argue that only by allowing them to fully internalize these spillover benefits will they have the optimal incentives to invest in creating, maintaining, and expanding their networks.

Our examination of spillovers in the IP context can shed significant light on this debate. There is little doubt that investment incentives matter. The ability to discriminate among end users and uses presumably would improve network owners' incentives by increasing expected returns to infrastructure investment. But the real questions are (1) how much would incentives improve? and (2) at what cost to society?

As we discussed in Part II.B, there is no reason to think that complete internalization of externalities is necessary to optimize investment incentives; at some point, there are decreasing returns in terms of improved incentives to allowing property owners to capture more of the value realized by users. Discrimination by network owners might increase network owners' returns and their incentives to build and maintain networks. But the magnitude of such increases—and the corresponding

---

providers.” Arshad Mohammed, Verizon Executive Calls for End to Google's 'Free Lunch,' Wash. Post, Feb. 7, 2006, at D1. According to AT&T CEO Edward Whitacre, Google and other companies “use my lines for free—and that's bull.” Spencer E. Ante & Roger O. Crockett, Rewired and Ready for Combat: SBC and Verizon Are Spending Billions to Stay Competitive in the Broadband Era, Bus. Wk., Nov. 7, 2005, at 110, 110. Whitacre states “I ain't going to let them do that . . . [T]here's going to have to be some mechanism for these people who use these pipes to pay for the portion they're using.” At SBC, It's All About “Scale and Scope,” Bus. Wk. Online, Nov. 7, 2005, at [http://www.businessweek.com/magazine/content/05\\_45/b3958092.htm](http://www.businessweek.com/magazine/content/05_45/b3958092.htm) (on file with the *Columbia Law Review*). Ironically, less than two years before, cable and telecommunications providers were steadfastly denying that they wanted to do any such thing. See Grant Gross, Advocates Push for Network Neutrality Policy, Network World, Apr. 5, 2004, at 33, 34 (quoting spokesman for cable and telecom industries describing network neutrality as “truly a solution in search of a problem”).

146. Christopher Yoo has argued against network neutrality because it would preclude network owners from vertically integrating and competing in the applications and content markets. He argues that vertical integration is important because it would enhance competition among last mile providers along a more diverse set of products and services. See Christopher S. Yoo, Beyond Network Neutrality, 19 Harv. J.L. & Tech. 1, 13–18 (2005). It is not altogether clear to us why he believes network neutrality precludes vertical integration. Network owners are fully capable of competing in applications and content markets under a network neutrality regime; they simply would not be allowed to discriminate in favor of their own products and services. See Frischmann & van Schewick, *supra* note 135 (manuscript at 20–24). If, on the other hand, his argument is that allowing network owners to internalize the full value of their networks will give them additional profits and so encourage them to spend those profits to build the network, we think Yoo's argument is merely an application of the Demsetzian internalization framework, and is wrong for the reasons we explained in the previous section.

benefits of increased internalization via discrimination by network owners—is not known or even knowable.<sup>147</sup> In the end, it is blind faith in private property rights (and frequently misplaced distrust of government) that drives the argument against network neutrality.<sup>148</sup>

More importantly, shifting from an Internet infrastructure commons to an Internet infrastructure for which access and use are allocated and prioritized according to users' willingness and ability to pay may entail significant social costs. Common nondiscriminatory access to the Internet infrastructure facilitates widespread end-user participation in a variety of socially valuable productive activities.

End-users . . . engage in innovation and creation; they speak about anything and everything; they maintain family connections and friendships; they debate, comment, and engage in political and nonpolitical discourse; they meet new people; they search, research, learn, and educate; and they build and sustain

147. It is worth noting as well that if in fact incentives to invest in infrastructure are suboptimal and need to be improved, there are alternative ways of doing so that do not entail discrimination on the basis of content or application. Some viable options include direct subsidization of infrastructure expansion, tax incentives to support the same, cooperative research and development projects, and joint ventures. In addition to improving incentives for private provision of infrastructure, government provision of infrastructure is another viable option for the last mile. In fact, municipal provision of broadband access to the Internet has gained significant momentum in recent years. But we are skeptical that any of this is necessary. Telecom companies radically *overbuilt* infrastructure in the late 1990s without any of the rights they now demand, and a number of companies are competing to build free public wireless networks in major cities throughout the country, again without any such guarantee.

148. The network neutrality debate is filled with arguments that falsely pit “the market” against “the government” in a battle for supremacy as the preferred system for allocating resources. Such arguments conveniently ignore, first, that the government has been, is, and will continue to be involved in the regulatory environment in one way or another; second, that the “market” for telecommunications infrastructure is not currently competitive and is becoming less so as telcos reassemble the Bell System via merger; and third, that managing infrastructural resources as commons and letting spillovers flow does not exclusively depend on either market or government allocation. An Internet infrastructure commons built on nondiscrimination rules may be an attractive allocation mechanism precisely because it avoids individualized allocation by either unregulated private or government decisionmakers.

Christopher Yoo, for example, argues that the government should not intervene and alter the status quo; government forbearance is the answer. Larry Lessig has also argued for maintaining the status quo, but he has argued that the end-to-end principle is what should be sustained. Yoo's baseline is the network owners' freedom from government regulation, and Lessig's baseline is users' freedom from network owners' discrimination. These conflicting claims of freedom give rise to the network neutrality debate but do not, in isolation, tell us how the debate should be resolved.

The “market versus government” argument is also particularly inappropriate in the telecommunications context because the market in question is not a competitive one, but one characterized by enormous economies of scale, long-term market power, and network effects. Whatever the merits of a free market against government regulation—and we think they are generally great—they have little bearing on a choice between government and an unregulated monopolist not subject to the discipline of market forces. See Lemley, *Ex Ante*, *supra* note 86, at 148–49 (making this point).

communities. . . . These are the types of productive activities that generate substantial social value, value that too easily evades observation or consideration within conventional economic transactions. When engaged in these activities, end-users are not passively consuming content delivered to them, nor are they producing content solely for controlled distribution on a pay-to-consume basis. Instead, end-users interact with each other to build, develop, produce, and distribute public and nonmarket goods. . . . [P]articipation in such activities results in external benefits that accrue to society as a whole (online and offline) [and] are not captured or necessarily even appreciated by the participants.<sup>149</sup>

Shifting to a system where access to and use of the Internet are allocated and prioritized according to users' willingness and ability to pay—which is the basic objective of network discrimination—preferences certain end-user activities (i.e., those that generate observable and appropriable benefits) over others (i.e., those that generate spillovers).<sup>150</sup> It is also likely to reduce innovation at the applications level, since more of the value of that innovation will be transferred to the owners of the network. And encouraging that applications-level innovation may be more important than encouraging additional innovation in the network itself.<sup>151</sup> In our view, the social opportunity costs of allowing network owners to dismantle the Internet's infrastructure commons may be tremendous but incredibly difficult to measure precisely because so much of the value generated by the Internet is not fully captured in market transactions. Preserving Internet spillovers requires preserving network neutrality.

#### IV. CONCLUSION: TOWARD A THEORY OF SPILLOVERS?

To this point, we have focused primarily on rethinking the value of spillovers and the economics of IP. Our reasoning, however, suggests some broader conclusions about the law and economics of spillovers. In this concluding Part, we explore a few tentative conclusions and areas for further study.

---

149. Frischmann, *Economic Theory*, supra note 3, at 1017–18 (footnotes omitted). For further discussion of such activities and the manner in which value is generated, see id. at 1018–20.

150. We recognize that network owners have the incentive to let some spillovers flow so as to increase the net value of their platform and encourage innovation in complementary markets for goods and services. See Farrell & Weiser, supra note 139, at 100–05. But there are limits to this. See id.; Frischmann & van Schewick, supra note 135; Barbara van Schewick, *Towards an Economic Framework for Network Neutrality Regulation*, 5 *J. on Telecomm. & High Tech. L.* (forthcoming 2007) (manuscript at 39–40, on file with the *Columbia Law Review*). While there is a risk of overt discrimination by network owners for anticompetitive reasons, we do not believe that the bias we are highlighting depends upon such behavior. As we discussed previously and in our other work, the bias persists even in competitive markets.

151. For this argument, see van Schewick, supra note 150 (manuscript at 38–40).

As we have discussed, the conventional law and economics thinking about externalities is that they are a bad thing, a market failure in need of correction. On this view, externalities interfere with rational decision-making and distort market allocation of resources. Internalization is needed, and private property rights are the preferred institutional mechanism for internalization.

But as our argument demonstrates, there are some fundamental flaws in this thinking.<sup>152</sup> First, externalities do not always distort market allocation. Many externalities are in fact irrelevant and thus not worth internalizing because internalization would not change anything other than the distribution of wealth. The marginal benefits of internalization approach zero at some point (for example, when there is no change in the producer's incentive structure). Second, there may be costs to internalization beyond the transactional costs associated with internalization itself. Internalizing spillovers may impact user behavior in a manner that reduces social welfare.

Together, these conclusions lead us to a simple but important point: Even where externalities distort market allocation, those distortions may be social welfare enhancing. Conversely, extending property rights to internalize externalities may distort market allocation in a manner detrimental to social welfare. We can generalize this point: Property rights distort resource allocation in the same way as externalities, although perhaps in the opposite, or at least a different, direction.<sup>153</sup> A Coasean analysis of externalities suggests that they are reciprocal in nature; externalities are neither positive nor negative in the abstract.<sup>154</sup> They are jointly produced by agents with interdependent relations, and whether an externality is positive or negative depends on how and to whom we allocate rights. Here we find a wonderful Coasean symmetry between externalities and property rights. When relevant, both externalities and property rights distort the market allocation of resources; when irrelevant, neither does. Under the same Coasean logic, as with an externality, a priori, one simply cannot say that the distortion caused by a new property right (or an incremental change in a property right) has a positive or negative effect on social welfare.

Coase warned us to avoid reflexively invoking externalities to justify government intervention in the form of taxation, regulation, or subsidies. Often, he showed, such intervention is unnecessary because affected parties will be able to work things out privately.<sup>155</sup> Our basic point is the mirror image: One cannot simply invoke externalities and the weakness

---

152. Portions of this and the next two paragraphs are adapted from Frischmann, Demsetzian Trend, *supra* note 6 (manuscript at 23–24).

153. Cf. Landes & Posner, *supra* note 102, at 58 (“Copyright, however, causes its own distortions in the market for expressive works.”).

154. See Guido Calabresi & A. Douglas Melamed, Property Rules, Liability Rules, and Inalienability: One View of the Cathedral, 85 Harv. L. Rev. 1089, 1102–05 (1972).

155. Coase, Social Cost, *supra* note 30, at 2–8.

of existing rights to justify propertization and privatization. The processes by which property rights are created almost always involve “government intervention” in one form or another. Such intervention may be unnecessary and in fact may lead to welfare-reducing distortions.

The importance of our argument extends well beyond property rights. Many laws, and perhaps even bodies of law, can be understood as attempts to internalize externalities. Environmental pollution is an archetypal example of an externality. Acme Factory produces widgets and in doing so emits pollutants into the environment. People living downstream or downwind from Acme receive the pollutants and bear some costs as a result. These costs are external to Acme’s decision to produce widgets unless the government forces Acme to take the costs into account by taxing Acme, for example. Environmental laws and regulation generally aim to force externality-producing agents, such as Acme, to fully account for the consequences of their actions.<sup>156</sup>

As we have discussed, however, it turns out that externalities are not always a problem in need of a solution. In some situations, externalities ought to be simply ignored. In those cases, the policy solution is to let the costs and benefits fall where they may and not to worry about whether parties have taken all costs or benefits into account. Internalizing so-called irrelevant externalities will not change behavior or resource allocation, so internalization is not worth the effort.<sup>157</sup> In such cases, there really is no problem for the law to solve.

More importantly, though, some externalities should be encouraged rather than eliminated or ignored. In many different contexts, the law actually is designed to encourage certain externality-producing activities. As we have discussed, IP laws unequivocally promote externality-producing activities—*invention, creative expression, and innovation*. Participation in these productive activities generates social returns that exceed private returns by a substantial margin.

Many public policy debates over legal and economic issues boil down to a debate over which types of externality-producing activities to be concerned with and the extent to which institutions should be designed to

---

156. Because Coase believed externalities to be reciprocal in nature, he disputed the notion that Acme *causes* homeowners to suffer a *negative* pollution externality; instead, he viewed the harm realized by homeowners as jointly produced by both Acme and the homeowners because they engage in interdependent activities—*manufacturing and homeownership*. See *id.* at 2; Cornes & Sandler, *supra* note 2, at 79–80, 86 (noting this); David de Meza, Coase Theorem, *in* *The New Palgrave Dictionary of Economics and the Law* 270, 273–74 (Peter Newman ed., 1998) (same). Coase brought attention to the possibility that the homeowners might avoid the harm by undertaking less costly preventative measures than those which might be available to Acme. This insight expanded the range of options for internalizing externalities. For a variety of reasons, however, government regulation remains the predominant mechanism for internalizing environmental externalities.

157. See Buchanan & Stubblebine, *supra* note 68, at 373–77.

regulate some and promote other externality-producing activities.<sup>158</sup> The broader lesson of our paper is that courts and scholars must resist the easy answer of equating public and private value by internalizing externalities. Spillovers aren't always bad, and more property rights aren't always good. Only if we understand when and why each can enhance social welfare can we hope to design legal rules that do more good than harm.

---

158. See Frischmann, *Economic Theory*, *supra* note 3, at 940–41.

