Regulating Competition in the Information Age: Computer Software as an Essential Facility under the Sherman Act

David McGowan

Follow this and additional works at: https://repository.uchastings.edu/hastings_comm_ent_law_journal

Part of the Communications Law Commons, Entertainment, Arts, and Sports Law Commons, and the Intellectual Property Law Commons

Recommended Citation
Available at: https://repository.uchastings.edu/hastings_comm_ent_law_journal/vol18/iss4/4

This Article is brought to you for free and open access by the Law Journals at UC Hastings Scholarship Repository. It has been accepted for inclusion in Hastings Communications and Entertainment Law Journal by an authorized editor of UC Hastings Scholarship Repository. For more information, please contact wangangela@uchastings.edu.
Regulating Competition in the Information Age: Computer Software as an Essential Facility Under the Sherman Act

by
DAVID McGOWAN*

Table of Contents

I. The Nature of the Problem: Competition, Synergy, and Strategic Behavior ................................................................. 778

II. The Curious History of the Essential Facilities Doctrine ... 781
   A. Natural Monopoly: *Terminal Railroad* and *Otter Tail* ............................................................................. 782
   B. Potential Doctrinal Expansion: *Aspen Skiing* and *Kodak* ................................................................................. 786
   C. The Proper Scope of the Essential Facilities Doctrine .................................................................................. 804

III. Networks, Externalities, and Innovation: Competitive Implications of Copyright Protection of Computer Software ................................................................................ 806
   A. Use of Copyright As A Barrier To Competition: ISO Copying Of Manufacturer Software .......................... 808
   B. Copying For Compatibility: Reverse Engineering, Replication, And The Essential Facilities Doctrine...... 816
      1. The *Sega* Decision And The Fair Use Defense .......... 817
      2. Reverse Engineering, Replication, And Essential Facilities ....................................................................... 833
   C. Strategic Contractual Behavior And The Essential Facilities Doctrine ......................................................... 836

* Associate, Howard, Rice, Nemerovski, Canady, Falk & Rabkin, San Francisco. The views expressed here do not necessarily reflect the views of Howard, Rice, its clients, or any other person. Annette L. Hurst of Howard, Rice and Mark A. Lemley of the University of Texas provided valuable comments and criticism on these topics; such errors as survived their scrutiny are the sole responsibility of the author.
D. Network Effects and the “Tipping” Argument for
the Essential Facilities Doctrine ........................................ 841

IV. Conclusion ........................................................................ 849
Introduction

Antitrust and copyright law both seek to enhance consumer welfare through various incentives and penalties. For antitrust, the goal of consumer welfare takes the form of an emphasis on firms’ pricing and output decisions; antitrust is generally satisfied if firms do not reduce output and price above marginal cost. Thus, for example, antitrust condemns conduct such as price fixing and horizontal market division in order to protect consumers from the deadweight losses such practices produce. Absent such protection, producers could extract supracompetitive profits from consumers by reducing output below the level that would be sustained by a competitive market, driving prices to a level higher than could be sustained by a competitive market. The antitrust laws thus give the price mechanism relatively free reign to achieve the optimal (efficient) allocation of resources.

Copyright’s conception of consumer welfare is slightly different; copyright seeks to promote innovation in and production of creative arts for the benefit of the public as a whole. Copyright pursues this goal through a rate of return analysis: borrowing from Samuel Johnson’s admonition that “[n]o man but a blockhead ever wrote, except for money,” copyright focuses on promoting investment in the creative arts by preserving for the creator (and his or her investors) such returns as the market will confer upon the exclusive control over copying of a work. The law reasonably fears that without such protection investment in the creative arts would be suboptimal because the creator would have to share the returns from copies of the work with others who had not borne any of the expense of creation. Works that would yield a sufficient return to be produced under a regime of copyright enforcement might not produce a sufficient return to justify production without such enforcement, and thus presumably would not be produced at all. Such returns may or may not be supracompetitive; as with other forms of property the returns will be a

---

1. 4 JAMES BOSWELL, LIFE OF JOHNSON 29 (A. Birrell ed., 1904).
2. See, e.g., Harper & Row Publishers, Inc. v. The Nation Enters., 471 U.S. 539, 546 (1985) (“The rights conferred by copyright are designed to assure contributors to the store of knowledge a fair return for their labors.”); William W. Fisher, III, Reconstructing the Fair Use Doctrine, 101 HARV. L. REV. 1661, 1687 (1988) (“[T]he elaborate combination of grants and reservations that comprise the Copyright Act is designed to advance the public welfare by rewarding creative intellectual effort sufficiently to encourage talented people to engage in it, while at the same time making the fruits of their genius accessible to as many people as possible as quickly and as cheaply as possible.”) [hereinafter Fisher, Fair Use].
function of market forces, most prominently the availability of acceptable substitutes for the copyrighted work and the ease with which such substitutes could be created. If close substitutes exist for a given item of intellectual property, or could readily be introduced in the short term, a grant of copyright will not confer the power to charge supracompetitive prices.  

For the most part, antitrust and copyright follow a parallel course, complementing one another in the promotion of consumer welfare by promoting investment in creative works and monitoring the structure of the markets in which such works compete. But conflicts, real or apparent, have begun to emerge in recent years, particularly in the software industry. Software firms (Microsoft being the obvious example) have begun to achieve a scale that, at least in the minds of some, equates with a monopoly position with respect to certain products or systems. Whatever their merit, these concerns may be addressed through traditional antitrust analysis and do not pose copyright issues in and of themselves. The situation may be different with respect to a number of more interesting issues that have arisen concerning the nature and scope of the rights that may be asserted in copyrighted software. These issues pose questions relating to how broad the protection of a copyright should be, and whether markets in computer software display economic characteristics—such as natural monopoly—that have long been considered sufficient to invoke antitrust and other forms of regulation.

Several firms have asserted in litigation that one or another type of software constituted an "essential facility" under the antitrust laws. For example, a firm that makes software providing additional features for a spreadsheet program might sue the maker of the program seeking a compulsory license on the ground that access to the software

3. See Fisher, Fair Use, supra note 2, at 1703; see also U.S. Dept. of Justice and Federal Trade Commission Antitrust Guidelines for the Licensing of Intellectual Property, § 2.2 (1995) ("[T]here will often be sufficient actual or close substitutes for intellectual property to prevent the owner from exercising market power.") [hereinafter DOJ Intellectual Property Guidelines].

4. Monopoly status, however, does not in and of itself violate the Sherman Act. A firm that achieves a monopoly simply by virtue of a superior product, or more efficient means of producing even a homogeneous product, does not violate the antitrust laws.

5. With respect to issues of firm size and supply and demand elasticities, intellectual property rights are simply assets, comparable with respect to these issues to assets any other sort of supposed monopolist might have. Though altering the scope of a copyright could certainly alter a firm’s market position, the same could be said for rights attendant to any property the firm used or sold. See DOJ Intellectual Property Guidelines, supra note 3, at § 2.2.
code, including such copying as was necessary to make the access meaningful, was essential to the continued viability of its own program. Or a firm that provides maintenance and repair services for certain machinery might sue the manufacturer claiming that it needed to run (and thus copy) certain software in order to service the machinery. As applied to Microsoft, the essential facilities notion is rapidly gaining popularity; even the *New York Times Magazine*, not previously known as a repository of antitrust theory, has published an article opining that Microsoft's architectures "are an essential facility—to use the antitrust jargon . . . ."6 More informed observers of the antitrust scene appear captivated by the idea. Robert Pitofsky, Chairman of the Federal Trade Commission, stated that "[t]he foremost question facing antitrust is when you get a firm that achieves monopoly and is also an essential facility, what should we do?"7

These issues highlight a point of divergence between the antitrust and copyright laws. Antitrust, to borrow a well-worn phrase, seeks to protect competition, not competitors.8 More formally, antitrust seeks to ensure that firms price at marginal cost rather than above or below that cost (though prices below marginal cost are worrisome only if they can later be recouped by imposition of prices above marginal cost).9 Although a debate on the subject continues, antitrust is not properly concerned with the distribution of wealth among firms, but only with wealth maximization through the efficient allocation of resources via the price mechanism. The economic analysis necessary to accomplish antitrust’s mission is taken largely from the fields of price theory and industrial organizations, which examine structural elements of markets and the operation of price as a method of allocating resources.

In contrast, copyright has legitimate firm-specific concerns. As noted above, production and innovation in intellectual property, as elsewhere, is a function of investment. Investments are functions of available returns, including returns that could be earned on

investments in fields other than intellectual property but with similar risk. These returns are partially a function of the strength of intellectual property rights. Strength in copyright is a function of two variables: the scope of copying that may be prohibited and the duration of the prohibition. Under present market conditions, software evolves very rapidly. Thus for most software the duration of a copyright (presently the shorter of 100 years from creation or 75 from publication, for works for hire) is essentially irrelevant past a fairly short term; copyrights could last 100 years or 10 without affecting the rate of return analysis very much. The scope of the copyright is more important. If the scope is too broad, net investment in intellectual property will decrease because the first firm to obtain a copyright will be able to sue later entrants for infringement. If the idea of a spy novel, for example, were itself copyrightable, there would be fewer spy novels because the copyright owner could either prevent entry or obtain a portion of the available returns for itself through licensing (and the owner would be in a very strong position to negotiate the license). Firms contemplating entry would have to account for the risk of infringement or the necessity for paying royalties; either factor would reduce the expected value of the investment, and thus the likelihood that investment would occur.

Extremely weak protection presents the same risks. If a firm's competitors can copy its software, and thus compete by offering substitutes without incurring research and development costs, then they will be able to undercut the firm's price and deny it a profit on its work. In such a regime, investment in intellectual property will be unattractive from an ex ante perspective. The risks of losing money will be high, and the expected value of an investment in intellectual property will therefore be low. This is a problem for both individual firms and the public, for if the intellectual property laws provide insufficient inducements for investment in that area, too little intellectual property will be produced and the public will be among the losers. The economic analysis necessary to accomplish copyright's mission should be taken from the field of financial economics, which examines, among other things, the relationship between risk and return and the proper method for evaluating whether firms should

10. This number is the opportunity cost a firm incurs by investing in intellectual property rather than the other field.
make investments. In reality this is almost never the case in published opinions and far too rarely the case in the academic literature. 12

In general one might say that antitrust’s mission is to see that the market operates freely, while copyright seeks to provoke a particular response from the market. Claims at the intersection of antitrust and copyright thus raise complex legal and economic issues, which have been addressed by only a few courts, constrained by the unclear contours of antitrust and the perhaps overly clear contours of the Copyright Act. From a policy perspective, the principal problem is to ascertain the mix of antitrust, copyright, and other legal principles that best enhances competition and innovation in software markets, and thus best promotes consumer welfare. This requires an understanding of the relationship between the scope of copyright protection and investment in the production of copyrighted works, as well as an understanding of the capacity and limitations of antitrust to efficiently monitor the exercise of economic power of any kind, including such power as is conferred by copyright.

This article discusses these issues and identifies the economic variables that courts should examine in evaluating antitrust essential facilities claims in the software context. This article concludes that antitrust in general, and the essential facilities doctrine in particular, is ill-suited to the task of policing the power of software copyrights. While many of the claims that have been brought raise legitimate legal and economic issues, they are primarily distributional issues relating to the optimal incentive structure for the promotion of investment in intellectual property. Because those issues must be resolved based upon the scope of power conferred by a copyright, they must be addressed by the copyright laws using the tools of financial economics; antitrust (and price theory) does not provide the analytical tools necessary to accomplish this task.

Part I of the article presents three situations in which the issue has arisen, describing the basic economic factors at work and the claims parties have made. Part II discusses the emergence of the essential facilities doctrine from the antitrust mists, and attempts to distill its animating principles from the few cases where such a claim has prevailed. Part III outlines the scope of a copyright grant in the software context, including the economic principles pertinent to these situations, and discusses the application of the essential facilities doctrine in software markets.

12. For a notable and welcome exception, See Fisher, Fair Use, supra note 2, at 1698-1705.
I

The Nature of the Problem:
Competition, Synergy, and Strategic Behavior

Most of the basic problems we are concerned with here can be seen by examining three hypothetical cases in which the elements of competition, synergy, and strategic behavior arise. The first case involves a dispute between parties (most likely competitors) who have no contractual relationship. One party requires access to copyrighted software code developed by the other in order to develop a product or provide a service. Cases of this nature include some reverse-engineering cases, in which software manufacturers seek access (through limited copying) to code written by their competitors in order to write programs compatible with their competitors' systems. Accolade's reverse engineering of Sega's code in order to write programs for games that would run on Sega's game consoles is an example of this type of case.13 Similar concerns arise when a firm seeks access to copyrighted material in order to provide some service, such as computer repair. Grumman's efforts to obtain copies (by means somewhat older and less refined than reverse engineering) of Data General's diagnostic software,14 and Peak Computer's similar efforts to make copies of MAI's software15 are examples of this variation on the problem.

The second type of case is characterized by a dispute between parties to an existing contractual licensing relationship, with attendant reciprocal rights, obligations, and expectations. For example, suppose a firm ("A") produces a home banking program. In order to enhance the value of its program, A licenses portions of its code to other firms (firms "C") that make software that allows consumers to customize the application program to their particular needs by adding features a customer wants such as calculating hedges, monitoring the Eurobond market, etc. This licensing arrangement carries obvious potential reciprocal benefits for the parties: the ability to purchase customized software will make the basic home banking program more attractive to consumers, while any growth in the market share of the basic program will make the license more valuable to C.

But while both programs in this hypothetical are more valuable together than they would be separately, the power structure of the relationship may be asymmetric, depending on consumer preferences and the costs of developing the different types of software. Suppose the basic application program was costly to develop, and consumers derived a large portion of the utility of the combined programs from the application program itself. In this case, A might have exploitable leverage over C, particularly if A's application program obtains a large market share and modifying C's program to work with competing programs (or of developing application software to support the customized program) is costly. C's cost of modifying its program to work with an application program with a smaller market share, as well as the opportunity cost of not selling in the larger market, will be referred to as C's "switching costs." Under such conditions, A may be able to extract benefits from C based in part on the value C has added to A's program. If A and C have a three-year license, A should be able to obtain fees and/or concessions in an amount slightly less than C's switching costs when negotiating an extension of the license. This example is not problematic, reflecting merely the risks inherent in any contractual relationship involving uncertainty over a given term. Problems may arise, however, if A seeks to renegotiate its license in mid-term.

Part of the leverage potentially available to parties in the second type of case stems from the economic phenomenon at the heart of the third case—network effects. A market is characterized by network effects if "the utility that a user derives from consumption of a good increases with the number of other agents consuming the good." A few common examples of network effects markets, discussed in greater detail below, should assuage any present economic anxiety brought on by the term: telephones are more valuable to existing owners when additional consumers buy phones because such purchases increase the number of people with whom each existing user may communicate.

16. The same problem could arise, for example, in the lease of a commercial property of uncertain value. Should the lessee establish a valuable business (say, an espresso bar next to a law school), it is exposed to a risk that its lease will not be renewed because the lessor wants to establish a similar business on the same site.

17. These two examples should be seen as points on a continuum ranging from competitive parties with no previous contractual relationship to parties with longstanding contractual ties in which one party engages in opportunistic behavior.

The same is true of fax machines, language itself, and many other things. Such markets may require a degree of technical standardization because fax machines and telephones may connect with each other only if each operates on the same standard.

Some fear that network effects will reduce social welfare by locking consumers into an inferior technology simply because so many people have chosen to use it—a phenomenon known as "tipping." Switching costs play a major role here, too, because existing users will incur costs (likely opportunity costs, but it doesn't matter) up to the point at which investment in new technology, both in terms of purchasing the technology and training employees how to use it, becomes profitable. The importance to antitrust is the fear that, if our hypothetical firm A succeeds in getting its software adopted as a standard, it might be resistant if not immune to competition from other standards due to the extent of its network. Entry will be more difficult because a large installed base of users is an inherent advantage of an existing system, regardless of its relative shortcomings when compared to other systems. Given the difficulty of dislodging a market standard, supply may be inelastic in network effects markets (relative to the elasticities that would apply to the technology without the network effect), and a given A firm may have some leeway to charge supracompetitive prices without fear of attracting entry. Other common fears are that a firm will: use power in a network effects market to leverage its way into other markets; charge predatory prices for some systems in order to lock consumers into a network effects aftermath; and rely on the network effects to recoup predatory losses.

These three hypothetical cases represent points on a continuum of problems concerning competition in markets involving intellectual property rights, network effects, and strategic behavior. There are many other combinations, and the analysis will depend on the characteristics of the market in each case. The intersection of antitrust and intellectual property is challenging, and the economic implications of network effects are not yet fully understood. Therefore, the task of reconciling the various interests involved to enhance social welfare requires a cautious, disinterested analysis rooted in the economic facts of each case. The theoretical framework necessary to interpret those facts requires a similarly disinterested analysis. Insofar as the essential facilities doctrine is concerned, developing such a framework is the subject of the following sections.
II

The Curious History of the Essential Facilities Doctrine

The late Professor Areeda termed the essential facilities doctrine "an epithet in need of limiting principles."\(^{19}\) Professor Hovenkamp concurs, casting the doctrine as "one of the most troublesome, incoherent and unmanageable bases for Sherman § 2 liability" and concluding that "[t]he antitrust world would almost certainly be a better place if it were jettisoned, with a little fine tuning of the general doctrine of the monopolist's refusal to deal to fill in the resulting gaps."\(^{20}\) Even these appraisals, coming from thoughtful commentators who could hardly be accused of slavish devotion to Chicago, may be too kind.

More than an epithet, the essential facilities doctrine embodies some of the most unfortunate tendencies of antitrust. Most prominent among these is antitrust's tendency to identify an economic state of affairs and, either wrongly perceiving the state of affairs to be a legitimate antitrust problem or wrongly perceiving it to be ubiquitous, to endorse broad principles far exceeding any justifiable remedial purpose, condemning innocuous or procompetitive practices in the process. The black letter of an essential facilities claim, such as it is, is best stated in *MCI Communications Corp. v. AT&T Co.*:

> The case law sets forth four elements necessary to establish liability under the essential facilities doctrine: (1) control of the essential facility by a monopolist; (2) a competitor's inability practically or reasonably to duplicate the essential facility; (3) the denial of the use of the facility to a competitor; and (4) the feasibility of providing the facility.\(^{21}\)

The most important of these elements is the second—whether a competitor may duplicate the "facility" in question.\(^{22}\) The doctrine has its origin in a case dealing with a natural monopoly and is out of place in any circumstance where the uncommon economic characteristics of such monopolies are not present.\(^{23}\) There is little if anything to recommend the doctrine as a tool for regulating competition in the software industry.

---

22. In most cases defining the facility is fairly easy; in software markets, however, more difficult questions arise. These are discussed in Part III.
We will examine four cases: *United States v. Terminal Railroad Association*, the case commonly cited as the origin of the doctrine; *Otter Tail Power Co. v. United States*, which represents a rare justifiable use of the doctrine; *Aspen Skiing Co. v. Aspen Highlands Skiing Corp.*; and, though it was not itself an essential facilities case, the Supreme Court's opinion in *Eastman Kodak Co. v. Image Technical Services, Inc.*, which opened the door to many of the essential facilities cases brought today.

A. Natural Monopoly: *Terminal Railroad and Otter Tail*

We begin with some rudimentary economics, which are necessary to understand the logic and limitations of the essential facilities doctrine, and also to understand the complex economics of the software industry. A firm's size is determined by the cost characteristics of its particular industry: in competitive markets firms will maintain the minimum efficient size that will capture all relevant scale economies. If it costs $1,000 to establish a widget-making firm, and each widget costs $1 to make, the cost of producing one widget is $1,001, which is what the firm must charge to recoup its costs; the cost of producing 1,000 widgets is $2,000 each. If the widget market can absorb one million widgets per year, and each widget firm can produce 100,000 widgets per year at a per-widget cost of $1, then ten widget firms will satisfy the market demand. If a widget firm could produce 500,000 widgets per year at the same per-widget cost, the market would have room for no more than two firms; the total cost per widget would be lower with two firms than with ten because the initial $1,000 cost would be spread across more units.

Similarly, if a firm could make one million widgets per year at the same per-widget cost, it would satisfy the entire market demand at the lowest cost, creating a "natural monopoly," a state in which a single firm can satisfy demand in the relevant market at a lower unit cost.

24. *Id.*
30. We here speak of economic cost, which includes a competitive return on investment, rather than out-of-pocket cost.
31. For a more detailed, and only slightly more technical, discussion of these points, see Hovenkamp, Antitrust Policy, *supra* note 9, at § 1.4a.
than two or more firms. Typical examples include local electric power and gas distribution, local telephone service, and, perhaps, long-distance transportation of petroleum and gas by pipeline. The problem arising from such markets is easy to state: "[I]f the market is occupied by a single firm, the firm will charge a monopoly price. If it is occupied by multiple firms, even assuming that the firms behave competitively, they will have higher costs and charge higher prices." As these examples bring to mind, natural monopolies are commonly controlled through governmental regulation of their prices, which is designed to provide firms in such markets with a competitive (rather than supracompetitive) return.

Because a natural monopoly market will by definition support only one producer, other producers cannot compete in the usual way—by attempting to provide substitute goods or services—because the market will not support two or more firms operating at minimum efficient size. This can create problems for firms that need to deal with natural monopoly markets. For example, almost all firms, from law firms to barber shops, use electricity as an element of production. But no firm could realistically integrate backwards into the production and/or distribution of electricity, nor induce entry into such markets, because the cost structure in those markets would not support more than one firm even if entry were otherwise feasible. Firms would thus potentially be at the mercy of the electric company, which is one reason why rates in such markets historically have been regulated.

The essential facilities doctrine is commonly held to originate with the Terminal Railroad case, which involved a group of railroads that joined together to purchase the only terminal facilities serving St.

---

33. Id.
34. See MCI Communications Corp. v. AT&T Co., 708 F.2d 1081 (7th Cir.), cert. denied, 464 U.S. 891 (1983).
35. Scherer, supra note 32.
36. See Hovenkamp, Antitrust Policy, supra note 9, at § 1.4b.
37. For example, the Federal Energy Regulatory Commission sets prices for interstate sales of electricity based on the return required by investors to hold securities with risks comparable to the risk of the utility's stock. This in turn requires a discounted cash flow analysis on a firm-by-firm basis. See Richard A. Brealey & Stewart C. Myers, Principles of Corporate Finance 54 (1991); Federal Power Comm’n v. Hope Natural Gas Co., 320 U.S. 591, 603 (1944) (“the returns to the equity owner [of a regulated business] should be commensurate with returns on investments in other enterprises having corresponding risks.”).
38. Areeda, supra note 19, at 842.
Louis. The Court apparently concluded in substance that the facility was a natural monopoly, and thus ordered that competitors of the members of the purchasing group be given access to the terminal on a nondiscriminatory basis. This order may have merely rearranged the deck chairs on the Titanic; it required the defendants to share their monopoly profits with their competitors, but did nothing directly to prevent the extraction of such profits from consumers. Of course, the Court's options were limited. It could not have undertaken the legislative task of regulating the terminal's returns, which would have required inquiries into the cost of capital and other, similarly complex and similarly empirical matters. It may be that forcing the defendants to share their monopoly rents with their competitors had some benefit, though (keeping in mind that in natural monopoly markets one firm can operate at lower cost than several) such an order could in theory simply have raised the costs of running (and thus using) the facility.

Over time, the Court added additional support for the essential facilities idea, most notably in Otter Tail, in which a firm that generated power (in that case a competitive market) also held a natural monopoly over the long-distance transmission of power (known as "wheeling") in the relevant geographic market. A few small cities in the market wished to sell electricity at retail to their residents, and thus asked Otter Tail either to sell power at wholesale to them or to wheel wholesale power generated by other firms, either

40. Areeda, supra note 19, at 842.
41. See Richard A. Posner, Antitrust Law: An Economic Perspective 208 (1976) ("It is difficult to understand how [the decree in Terminal Railroad] protects the public; its purpose and effect are, rather, to let the defendants' competitors share in the monopoly position enjoyed by the defendants.").
42. Professor Areeda appears to defend the opinion on this ground, stating that "the Supreme Court wisely concluded that the most efficient remedy was to admit nonmember competitors to the consortium." Areeda, supra note 19, at 842. This conclusion is not inescapable. To the extent that consumers were investors, for example, they might have recouped some of their losses in the form of an increase in the plaintiffs' stock price. As argued in more detail below, modern portfolio theory and the ability of consumers inexpensively to diversify their holdings suggests that the allocation of wealth, including monopoly rents, will even out over time and among firms as a whole, lessening the force of any arguments for the use of antitrust as a distributive vehicle. Such considerations may well have had less force in 1912, when capital markets were arguably less efficient and diversification more difficult; nevertheless, it does not follow that a distributional remedy was appropriate.
43. Otter Tail, 410 U.S. at 366.
44. Id. at 368.
45. Id. at 370.
option would have increased competition for consumers in the retail power market. Otter Tail, whose prices for wheeling were subject to federal regulation, used its natural monopoly over the power lines to ensure an effective monopoly in the unregulated retail market.\textsuperscript{46} Though the Court did not invoke the essential facilities doctrine as such, it held that the refusal to sell power to the cities or wheel power from other firms violated section 2 of the Sherman Act.\textsuperscript{47}

Four aspects of \textit{Otter Tail} are worth noting. First, it expanded the essential facilities concept from coordinated conduct among many firms, which characterized \textit{Terminal Railroad}, to the unilateral activity of a single firm. This is important because coordinated conduct among firms poses greater dangers of anticompetitive conduct, and a lower risk of anticompetitive rulings from the courts, than does conduct by a single firm. Because single-firm conduct, including refusals to deal with competitors, is ubiquitous, the implications of applying the essential facilities theory to such conduct are far greater than in the \textit{Terminal Railroad} context.\textsuperscript{48} Second, Otter Tail's power transmission lines were a natural monopoly, and the case thus does not expand the scope of \textit{Terminal Railroad} with respect to issues of market power or supply elasticity.\textsuperscript{49} Third, Otter Tail was subject to extensive federal regulation of its dealings, including its prices, with other firms;\textsuperscript{50} the Court's order that it was required to deal with them thus shifted the debate into a forum in which rates of return could (at least in theory) be regulated.\textsuperscript{51} Fourth, Otter Tail attempted to use a natural monopoly over one market (wheeling) to retain a monopoly (but not a natural monopoly) in another (retail sales of power).\textsuperscript{52} The Court's opinion thus condemned the use of natural monopoly power to create what Judge Easterbrook has aptly called "bottlenecks" in competitive processes.\textsuperscript{53}

\textsuperscript{46} \textit{Id.} at 371.
\textsuperscript{47} \textit{Id.} at 377.
\textsuperscript{48} Professor Areeda discusses this point in greater detail in Areeda, \textit{supra} note 19, at 849.
\textsuperscript{49} \textit{Otter Tail}, 410 U.S. at 369.
\textsuperscript{50} \textit{Id.} at 371.
\textsuperscript{51} For a discussion of these aspects of the \textit{Otter Tail} opinion, see Areeda, \textit{supra} note 19, at 848; Fishman v. Estate of Wirtz, 807 F.2d 520, 571 (7th Cir. 1986) (Easterbrook, J., dissenting).
\textsuperscript{52} \textit{Otter Tail}, 410 U.S. at 368-71.
\textsuperscript{53} \textit{Fishman}, 807 F.2d at 573 (Easterbrook, J., dissenting).
B. Potential Doctrinal Expansion: Aspen Skiing and Kodak

Although both Terminal Railroad and Otter Tail were natural monopoly cases, to some members of the antitrust bar they simply invited firms to demand, with occasional success, access to assets owned by the firm’s competitors to compete more effectively. Such enticing fruit was too much for some firms to resist, resulting in a crop of odd claims aptly summarized by Professor Areeda as including:

- a rock impresario seeking admission to the local auditorium;
- a teletype machine marketer complaining that its competitor will not sell machines for it;
- a ski resort complaining that a rival will not engage in joint marketing with it;
- a maker of “muscle building” food supplements demanding that a bodybuilding magazine accept its ads;
- a paper retailer complaining that other paper retailers will not admit it to their wholesale buying co-op;
- an anesthesiologist insisting that the local hospital, using in-house anesthesiologists, allow him to perform anesthesiological services as well; or
- the would-be oil seller, who has no storage tanks of his own, demanding to use those of an incumbent seller—to say nothing of Berky, who wants to know the results of Kodak's research before Kodak markets its own innovations.54

To this list we may now add: Alaska Airlines, which wants less expensive access to reservations systems run by American and United,55 TV Communications Network, which views ESPN and TNT as essential to its efforts to sell cable television subscriptions in Denver;56 and Dr. Tarabishi, whose staff privileges at the McAlester Regional Hospital are equally essential.57

Anecdote is no substitute for analysis. Still, it is difficult to resist the conclusion that those aspects of the doctrine which spawned this lot of cases should have been discarded long ago. Antitrust has a way of preserving such theories, particularly when they are malleable enough to fit a variety of sympathetic factual scenarios. This may explain the unfortunate appearance of Aspen Skiing Co. v. Aspen Highlands Skiing Corp.58 in the essential facilities line and, quite improbably, in software essential facilities claims. In these instances

54. Areeda, supra note 19, at 843-44 (citations omitted).
57. Tarabishi v. McAlester Reg’l Hosp., 951 F.2d 1558 (10th Cir. 1991).
the opinion has been said by plaintiffs to justify the use of the essential facilities doctrine to compel cooperation by a monopolist with a competitor.\(^59\) From the perspective of the Court's actual opinion, there are two problems with such invocation: the Court did not address the essential facilities rationale, and the defendant was almost certainly not a monopolist.

Defendant Aspen Skiing Company owned three of four ski mountains in Aspen Colorado; plaintiff Aspen Highlands owned the fourth.\(^60\) For 15 years the two companies had offered some form of cooperative ski ticket, which allowed consumers to ski at the mountains of either firm by buying the same ticket.\(^61\) Consumers preferred such packages, which gave them daily flexibility in choosing mountains to ski without having to purchase tickets (and stand in lines) each day.\(^62\) Aspen then ceased its cooperation and refused to offer joint tickets.\(^63\) This did not mortally wound Aspen Highland, though it felt relegated to "a day ski area in a destination resort"\(^64\) and lost some revenue.\(^65\) Nevertheless, Aspen Highlands sued, claiming that Aspen was a monopolist and that (i) its mountains (in the form of Aspen's participation in the joint ticket program) were an essential facility, without access to which Highland's competitiveness was diminished, and (ii) Aspen's refusal to cooperate with Highlands in offering the joint ticket was willful maintenance of monopoly power, in violation of section 2 of the Sherman Act.\(^66\)

The Court did not reach the essential facilities issue because it found ample evidence to support the jury's verdict on the plaintiff's section 2 claim.\(^67\) The section 2 claim in turn was lost largely because Aspen did not challenge the jury's finding that it was a monopolist.\(^68\) This was a mistake: Aspen was a "destination resort,"\(^69\) meaning its customers came long distances to ski there; they could just as easily have gone to Snowbird, Steamboat Springs, Jackson Hole, or any of a

\(^{60}\) Aspen Skiing, 472 U.S. at 593.
\(^{61}\) Id. at 589-93.
\(^{62}\) Id. at 589.
\(^{63}\) Id. at 593.
\(^{64}\) Id. at 594 (citation omitted).
\(^{65}\) Id.
\(^{66}\) Aspen Skiing, 472 U.S. at 595-96.
\(^{67}\) Id. at 611 n.44.
\(^{68}\) Id. at 596.
\(^{69}\) Id. at 587.
number of resorts in the Rocky Mountains. Aspen competed with such resorts regardless of its arrangement with Highland. It is thus highly improbable that Aspen had a monopoly in any properly defined market. This fact made writing a monopolization opinion difficult, and severely distorted the usefulness of the opinion in future cases.

The worst problem was that the alleged wrongful conduct was Aspen’s refusal to cooperate with its competitor for the competitor’s benefit. Although it would be hard to design a study to provide confirmation, both theory and intuition suggest that such conduct is almost always procompetitive. Among other things, such refusals allow firms to preserve the returns of their investments for their investors; the ability to exclude those who did not invest from sharing in such returns leaves more for the actual investors and thus increases the likelihood of investment in general. This is a key concept in copyright as well. As noted above, the copyright laws are designed to promote innovation, and innovation requires investment, which requires that investors be assured a competitive return. Hence the appearance of Aspen Skiing in copyright disputes is a dangerous development from both the antitrust and the copyright perspectives.

Because Aspen was not a monopolist, and the behavior at issue is generally procompetitive, it was difficult for the Court to write an opinion that did not provide significant comfort to firms looking to file antitrust suits attacking their competitors’ procompetitive conduct. The Court’s opinion, which stresses Aspen’s history of cooperation with Highlands, and sustains a jury verdict based on the cessation of such cooperation, does just that. The Court’s emphasis on cooperation was unfortunate. Counsel advising clients on antitrust issues have little choice but to warn that once cooperation is begun antitrust may be invoked as, in Professor Areeda’s felicitous words, “an antidivorce statute.” Thus, in support of its antitrust counterclaim to DG’s

70. The problem originated in the trial court, which instructed the jury that it could find both a relevant market and a relevant sub-market. “Thus instructed, the jury found that the market was downhill skiing at North American destination ski resorts and that the submarket was downhill skiing at Aspen. This ignores the critical issue: Which market did the defendant need to control in order to be able to successfully charge monopoly prices?” Philip Areeda, Monopolization, Mergers, and Markets: A Century Past and the Future, 75 CAL. L. REV. 959, 979-80 (1987).


72. Areeda, supra note 19, at 850.
copyright action in *Data General*, Grumman could plausibly cite *Aspen Skiing* for the contention that a "monopolist that has helped a market to develop may not withdraw its support without legitimate business justification." All this is troubling because there are many areas, particularly research and development, in which cooperation among firms is desirable. At some point, most cooperative undertakings will cease, and the firms will go their separate ways. Such cooperation (and its termination) is generally structured through contracts, most often express but possibly implied through conduct. Contract, rather than antitrust, would seem to be the proper repository of any antidivorce principles. If the parties' relationship did not amount to an implied agreement, it is hard to see why antitrust should be invoked.

Had Aspen actually been a monopolist, the Court's opinion might have been of some use. As it is, one is left with the conviction that antitrust would have been better off if the Court had used its power to dismiss the grant of certiorari as improvident—little if anything useful would have been lost and much confusion avoided. This is particularly so in light of the Court's later opinion in *Eastman Kodak Co. v. Image Technical Services.* This case is perhaps most responsible for the movement to infuse intellectual property disputes with antitrust principles. *Kodak* is worthy of extended discussion here for two reasons. First, the opinion held that "aftermarkets" (markets for goods or services related to only one product) were legitimate markets in which to allege an antitrust violation. Second, the Court articulated a theory that is significant for intellectual property cases and has been

73. *Data General*, 36 F.3d at 1188.

74. In particular, true monopoly power combined with the possible exclusionary effect of refusing to recognize tickets sold by Highlands, which Professor Areeda considers a possible form of exclusive dealing, might have justified the Court's result. Areeda, *supra* note 19, at 850. Absent monopoly power, this theory has little appeal. From the practitioner's perspective, one can only pray the Court to write no more opinions based on an assumption of monopoly power where no such power has been shown to exist, regardless of the mistakes of counsel on appeal. Law operates by analogy, albeit imperfectly, and sometimes by extremely bad analogy. Nevertheless, when counsel has a client who is not a monopolist, and is cited to a monopoly opinion that an opponent argues is analogous to the case at hand, counsel should be able to identify for the court all the competitive constraints operating upon the client that were ineffective in the case of the monopolist. This requires a comparison of facts. The problem, of course, is that if the firm in the cited precedent was not a monopolist it may well have been subject to constraints similar to the ones operating upon the client, and economically valid points of distinction will be legal failures.


76. *Id.* at 481.
aptly described by Professor Shapiro as “installed based opportunism.”

Plaintiffs in *Kodak* were independent service organizations (ISOs) that provided service for copiers manufactured by Kodak. They were independent because they were not affiliated with Kodak, which offered service for its copiers itself. The ISOs charged less for service than Kodak, which responded by implementing a policy of selling replacement parts only to consumers who purchased its service as well as its machines. Kodak also persuaded the manufacturers of these parts not to sell parts that fit Kodak machines to anyone other than Kodak. This made it difficult if not impossible for ISOs to obtain parts and thus to service Kodak machines. Some ISOs were forced out of business, and some lost significant revenues.

The ISOs sued, claiming that there were three markets involved in the sale of Kodak copiers: the primary market for the copiers themselves, in which Kodak competed with other manufacturers; a market for parts for Kodak copiers; and a market for servicing Kodak copiers. As the case stood before the Court, the ISOs had conceded that the primary market was competitive. Nevertheless, the ISOs claimed that Kodak had a monopoly in the market for parts for its copiers, and that it had tied service for Kodak copiers to sales of Kodak parts. The parts monopoly presumably existed because Kodak parts were not interchangeable with its competitors’ parts; only parts that fit Kodak machines could be used to repair Kodak

---

77. Carl Shapiro, *Aftermarkets and Consumer Welfare: Making Sense of Kodak*, 63 *Antitrust* L.J. 483, 487 (1995) (“The equipment manufacturer can engage in ex post exploitation of locked in buyers, also called installed-based opportunism, by making unexpected changes in aftermarket policies that exclude aftermarket rivals and thus allow the manufacturer to extract more money from locked-in buyers who have already purchased its equipment.”) [herinafter Shapiro, *Aftermarkets*].
78. *Kodak*, 504 U.S. at 455.
79. Id. at 457.
80. Id. at 458.
81. Id.
82. Id.
83. Id.
84. Id. at 457. The Court agreed that three such markets existed, relying on a test articulated in Jefferson Parish Hospital District No. 2 v. Hyde, 466 U.S. 2, 21-22 (1984), under which separate markets exist if there is sufficient consumer demand for the products to be sold separately. This test is unobjectionable in and of itself, but one must apply it carefully, with an eye toward the consumer welfare goals of the antitrust laws.
86. Id. at 459.
machines. This "monopoly" was created by Kodak's development of the machines, including their component parts. Kodak could manufacture the parts, or provide the necessary specifications to manufacturing firms, as efficiency dictated. As Justice Scalia pointed out in dissent, in this sense, Kodak's monopoly over parts was the same as the monopoly of every durable goods manufacturer over parts for its machines. Volvo has the same monopoly over its crankshafts. This sensible point provoked a disturbing footnote from the majority, which invoked the leverage theory of tying in response, citing a series of erroneous tying opinions with approval to magnify the damage done. Equally troubling, the majority indulged in the use of market share as a proxy for monopoly power, finding that "Kodak controls nearly 100% of the parts market . . . with no readily available substitutes . . . ." The insight that the owner of a Kodak copier has

87. Id. at 482.

88. Id. at 489.

89. As Professor Hovenkamp said, "[K]odak, like most other manufacturers of durable goods subject to repair, has a great deal of control over the manufacturing of its own repair parts and . . . its own repair parts are the only ones that will fit in its machines." Herbert Hovenkamp, Market Power Aftermarkets: Antitrust Policy and the Kodak Case, 40 U.C.L.A. L. Rev. 1447, 1454-55 (1993) [hereinafter Hovenkamp, Aftermarkets].

90. Kodak, 504 U.S. at 479 n.29. The Chicago School has long argued that a monopolist has a fixed quantum of monopoly power, which cannot be magnified by allocating it between markets through a tie. See, e.g., Robert H. Bork, The Antitrust Paradox: A Policy At War With Itself 372-75 (1993); see also Hovenkamp, Antitrust Policy, supra note 9, § 7.6 (arguing that leverage theory of tying is dubious basis for antitrust).

Professor Kaplow argues that it may be possible for a monopolist to increase the profits of its power by leveraging itself into a more profitable market—for example one characterized by network effects or protected by intellectual property rights. See Louis Kaplow, Extension of Monopoly Power Through Leveraging, 85 Colum. L. Rev. 515, 515-16 (1985). This is less of a concern in connection with aftermarkets because any social welfare loss in an aftermarket is strictly constrained by the primary market share of the aftermarket monopolist. If the primary market is competitive, as it was conceded to be in Kodak, leveraging does not present a coherent theory of social welfare loss. At the aftermarket level, leveraging concerns are primarily distributional, and thus present a weaker case for invoking antitrust.

91. Kodak, 504 U.S. at 479-80 n.29. The opinions on which the majority relied include International Salt Co. v. United States, 332 U.S. 392, 402 (1947) (affirming summary judgment against a defendant leasing salt processing machines with a requirement that the lessee also buy salt, without inquiring whether the defendant held market power in any market), and Northern Pacific Railroad Co. v. United States, 356 U.S. 1, 12 (1958) (affirming judgment against a defendant that sold land along its right of way on condition that the purchasers ship over its lines without considering that the defendant could not have had power in the market for land for manufacturing purposes).

92. Kodak, 504 U.S. at 481. The endorsement of market share as a proxy is unfortunate because even a firm with 100% of a market may not have market power, depending on elasticities of supply and demand. See, e.g., Los Angeles Land Co. v. Brunswick Corp., 6 F.3d 1422, 1425-26 (9th Cir.), cert. denied, 510 U.S. 1197 (1993).
no readily available substitutes for parts that fit Kodak copiers does little to advance any meaningful conception of antitrust policy. As Professor Hovenkamp put it, such factors “are, by themselves, absolutely irrelevant to the manufacturer’s power in aftermarkets.”

By focusing on parts and service as such, the Court adopted the ISOs’ perspective, which is to say the perspective of a competitor, rather than the perspective of the actual consumers at issue—purchasers of Kodak copiers. As a result, the opinion fails to come to grips with the fundamental fact that consumers had no desire for Kodak parts or service as such; they had a desire for copies. Parts and service are demanded only to the extent necessary to keep the copier in good enough repair to produce copies. Thus, assuming for the moment that Kodak charged “supracompetitive” prices for service, consumers would view such prices as part of the overall cost of obtaining copies, not in isolation. If Kodak service cost $25,000 more than ISO service over the life of the copier, which itself cost $500,000, the overcharge to the consumer is 5%. This may or may not be a material sum to the purchaser; it may mean life or death to the ISO. Nevertheless, the 5% overcharge would be all that is of concern to antitrust because that is all that is relevant to consumers. As importantly, this approach focuses on the market in which Kodak was forced to live: the primary, competitive market for copiers. There is little demand for either service or parts for Kodak copiers if Kodak’s fortunes in this market sag.

In response to Kodak’s argument that the competitive primary market prevented it from exercising market power in the aftermarkets for parts and service because consumers would then purchase its competitors’ machines, the Court seized on what it assumed were “market imperfections” in the primary market: switching costs and information costs. As this is the aspect of Kodak analogous to the software debate, the Court’s analysis deserves close attention.

The Court’s switching costs rationale was simple: existing owners of Kodak copiers would have to, among other things, incur some expense to dispose of those copiers, obtain new ones from a different manufacturer, integrate the new system into their business, and retrain
staff. Kodak could theoretically charge some amount slightly below that cost, but well above its marginal cost, for service, thus earning supracompetitive revenues (assuming Kodak's marginal costs were comparable to the ISO costs) without losing any customers. This point is correct but limited. Kodak could only earn such revenues from existing customers, and in particular those whose machines were not near the end of their useful lives (and thus in the market for a new copier). As Kodak rightly argued, however, doing so would be a "short-run game." Future purchasers would likely learn of its opportunistic behavior (though the rate at which this information would be disseminated is an empirical matter that could vary among markets) and take steps to prevent it, including buying from other firms in the competitive primary market, thereby limiting Kodak's "power" to extract supracompetitive profits to the life of one machine. This, however, is not a complete response to the switching costs argument. If nothing else, when Kodak eventually decides to exit the market there will be an end-period problem in which the company's concerns for its reputation will act as less of a constraint on extracting revenues from its installed base.

Although these factors do not provide conclusive answers for all cases, they would likely ameliorate the harm to consumers in most situations. And why the antitrust laws should be invoked to prevent what is essentially contractual opportunism over the life of a single product is not readily apparent, nor is any persuasive answer likely to exist. As discussed below, contractual provisions might well take care of a significant portion of the concern. Even if contractual provisions could provide only imperfect protection, it is still not clear why the antitrust laws should be applied to such a narrow problem. Perhaps because of the limited scope of the switching costs argument, the Court offered an additional theory of market failure, based on

96. Id. at 476.
97. This was a point of some debate. If, as Kodak contended, the ISOs were free-riding on its investment in copiers, the ISOs presumably had a lower cost structure. In that case, Kodak's higher rates would not necessarily reflect supracompetitive pricing. See id. at 483-86; see also Hovenkamp, Aftermarkets, supra note 89, at 1455 n.32.
98. Kodak, 504 U.S. at 473.
99. Id. at 470.
100. The constraint is not completely lost. If Kodak's copiers are replaced by a new generation of technology that Kodak also manufactures and if (as we may reasonably assume if the new generation technology is a substitute for the old) installed-base owners of Kodak copiers will be prospective purchasers of the new technology, Kodak will still face a significant reputational constraint. See Shapiro, Aftermarkets, supra note 77, at 490.
supposedly prohibitive information costs. This was the crux of the Court's reasoning from an antitrust policy perspective because it had the potential to transform Kodak's policy from a short-run strategy of contractual opportunism into a profitable general business practice undisciplined by well-informed consumers purchasing in a competitive market. The Court's theory was that "[l]ifecycle pricing of complex, durable equipment is difficult and costly. In order to arrive at an accurate price, a consumer must acquire a substantial amount of raw data and undertake sophisticated analysis." The Court argued that many consumers would be unwilling to incur such costs, and thus would be vulnerable to ex post opportunism by Kodak.

The lifecycle pricing idea was less of an insight than it might seem. Large purchases inevitably involve complex cost calculations, whether the consumer is aware of them or not. If payment is made over time, interest expenses must be considered and the best option chosen from available financing alternatives. Amortization and liquidation value must be considered, which require consideration of secondary or scrap markets and the selection of a terminal point for the investment.

Pricing the total cost of obtaining copies, which includes copier parts and service, is no different. A rational consumer will seek such information until the marginal expense incurred in obtaining it equals any cost savings realized. A consumer may rationally choose not to learn all the details of lifecycle pricing, which would in many cases imply that the costs at issue were not material relative to the overall cost of whatever system was at issue, and will factor any resulting uncertainty into the decision to purchase. Either way, the lack of perfect information in no way implies a market imperfection of a sufficient magnitude to warrant the invocation of antitrust. One of the main points of the Coase theorem is that information is quite often imperfect; while this point implies that legal entitlements should be allocated in such a way as to minimize the losses from bargains too costly to strike, it does not even begin to follow that antitrust is remotely capable of providing such a remedy, much less that it should be applied in such cases. Many consumers fail to read the agreements they sign for durable goods, and thus may find themselves victims of unwelcome surprises at a later date. But this cannot plausibly serve as

101. _Kodak_, 504 U.S. at 474.
102. _Id._ at 473.
103. _Id._ at 474-75.
a warrant for invoking antitrust. Not surprisingly, even defenders of the Court’s opinion find its information cost theory less than compelling. Professor Fox, whose defense we will examine in a moment, concedes that “[s]mall information failures or lags are garden variety disorders of economic life. Consumers must take responsibility to get information for themselves when the information matters to them.”

As Kodak pointed out, its competitors would likely be more than willing to provide consumers with such information if Kodak’s prices in the aftermarkets were too high. The Court wrongly discounted this probability, but it is a detail either way. Even if consumers do not have accurate price information, the conclusion that antitrust should be invoked requires a further policy analysis the Court never undertook.

106. Kodak, 504 U.S. at 474.
107. The Court believed that
A competitor in the equipment market may not have reliable information about the lifecycle costs of complex equipment it does not service or the needs of customers it does not serve. Even if competitors had the relevant information, it is not clear that their interests would be advanced by providing such information to consumers. Id. (citation omitted). The former point is unlikely; a prudent purchaser could look to the sources consumers look to every day for information on large purchases: current and former owners of the equipment, trade publications, repeat players such as lenders who finance purchases, and similar sources. See Bork, supra note 90, at 438 n.*. In fact, Kodak’s problems with its aftermarkets were sufficiently well-known to the industry for “[t]wo of the largest consumers of service and parts” to submit briefs to the Court “contend[ing] that they are worse off when the equipment manufacturer also controls service and parts.” Kodak, 504 U.S. at 479, n.28. That being the case, one would expect consumers to address any concerns at the time of purchase, an option available to all but those who owned machines at the time Kodak’s policy changed. As discussed below, those are the only purchasers for whom the Court’s concerns had any validity, and their plight, though real, seems insufficient to warrant the intervention of antitrust. The Court’s latter point, that competitors might not provide information on lifecycle costs, Kodak, 504 U.S. at 474, is perplexing. The Court’s basis for the statement was its concern that “[t]o inform consumers about Kodak, the competitor must be willing to forgo the opportunity to reap supracompetitive prices in its own service and parts markets.” Kodak, 504 U.S. at 474, n.21. This of course simply ignores the plaintiffs’ concession that the primary market was competitive. The Court is positing oligopolistic behavior, not a competitive market. See Bork, supra note 90, at 438 n.*. The Court was also worried that competitors would stay silent because informing customers might lead Kodak to lower its prices, thus making its copiers more competitive. This suggestion seems implausible at best. The Court posits a competitor willing to stand mute while a potential customer decides between its products and those of a competitor for fear that trying to make the particular sale at hand will lessen sales in the future. By parity of reasoning, one might expect the firm to send all its customers to a competitor so as to fool the competitor into thinking it was a monopolist. Once the duped faux monopolist raised its prices to monopoly levels, the firm could then profitably undercut its competitor’s prices.
We may begin this analysis by examining the effect of Kodak's policy, taking everything the ISOs alleged to be true. On such facts, owners of Kodak machines pay more to Kodak for service than they would in a competitive market, resulting in a material transfer of wealth from copier consumers to Kodak and the elimination of ISOs from the market. We may not infer, however, that wealth is transferred in sufficient amounts to force consumers to discontinue use of their copiers. Even as a monopolist, Kodak would likely not go so far. As noted above, consumers' use of copiers, and whatever such use adds to social welfare as a whole, is a function of the total cost of obtaining copies, not of the costs of service or parts, which may be a fairly small portion of the total cost even if that small portion is all or nothing to the ISOs. It follows that the real concern in Kodak is not so much with the deadweight loss that is the evil with which antitrust is traditionally concerned, as it is with the transfer of wealth from the copier owners to Kodak.

In this respect, Kodak poses a classic problem, more familiarly posed by the hypothetical of perfect price discrimination, of whether antitrust is concerned with pure (or close to pure) wealth transfers.

Assuming the ISOs establish that Kodak's policy successfully transferred wealth from Kodak's customers to Kodak, and assuming little or no reduction in output, it is legitimate to question why this should be an antitrust problem at all, particularly in light of the

---

108. Instead, Kodak would charge an amount slightly less than the maximum price a user would be willing to pay while continuing to use the machine.

109. In normal circumstances, monopolists by definition "will operate in the elastic portion of their demand curves." See Posner, supra note 41, at 10. It is possible that by increasing the price of service Kodak would induce customers to purchase less in the parts or service aftermarkets by, for example, waiting longer between service calls. This is not a restriction on output in the traditional sense. So long as Kodak sets prices to maximize its monopoly revenues, it will not price high enough to induce consumers to leave their machines broken rather than having them fixed. To the extent either parts or service costs are variable with the number of copies made, monopoly pricing in the aftermarkets might induce substitution toward less expensive copying methods. Depending on the viability of such substitution (cross-elasticity of Kodak's copiers with available substitutes), this could be a legitimate theoretical concern if the problem was systemic; that is to say, if the Court's claims of market failures due to information costs were a persuasive warrant for invoking the antitrust laws. Because they are not, we again deal with a problem limited to one group of owners for the maximum life of one copier. If there are few substitutes for Kodak's machines, as the Court's opinion implies, Kodak, 504 U.S. at 456-57 n.1, there is likely to be little if any substitution, and thus little if any social welfare loss.

concession that the primary market was competitive. Consumers are, among other things, investors. Diversified investors, which is to say most investors, should be indifferent to the allocation of the sums at issue. Over time and in the economy as a whole, winners and losers in such allocative debates even out. Employees, who are also investors through pension plans and the like, would have greater firm-specific interests, but as a general matter it would be impossible to know \textit{ex ante} which side of the dispute a given firm would fall on, and thus impossible to give effect to this interest. Kodak’s employees would want their company to win, while the ISOs’ employees would want the ISOs to win. To the extent either employees or consumers are investors, however, they would prefer the choice of the rule that maximizes wealth as a whole, regardless of distribution. These considerations, along with the likelihood that any deadweight loss in Kodak would be small, weigh against invoking antitrust as a redistributonal tool.

The ISOs, of course, would take strong issue with this, insisting that whatever abstract arguments one might choose to make about the distribution of wealth and wealth transfers in general, the issue is one of life and death to them, and their interest should not be ignored. Without competitors, they would argue, there can be no competition. Such arguments underlie Professor Fox’s defense of Kodak; as she sees it, Kodak “is a chink in the armor of the allocative efficiency model. Through the slit in the armor shines the light of legitimacy and justice: the right of well-performing firms, valued by consumers, not to be cut out of markets by a firm with power.”

Professor Fox believes this right is conferred by the Sherman Act and is the true source of the opinion, rather than, as some have argued, a sort of “post-Chicago” economic view of antitrust. From this

\begin{itemize}
  \item 111. This is because, from an investor’s perspective, proper diversification eliminates firm-specific risk; such investors are concerned about the returns from a market portfolio of assets, and thus favor rules that create wealth to rules that sacrifice some wealth for broader distribution of the remainder. \textit{See, e.g.,} \textit{Brealey \& Myers, supra} note 37, at 139 (discussing effects of diversification). Similar considerations suggest that bondholders and other corporate creditors, who presumably diversify their portfolios as well, would also prefer the wealth-maximizing rule.
  \item 113. Fox, \textit{supra} note 105, at 760.
  \item 114. Chicago’s demise has in any event been greatly exaggerated. The analytical approach associated with Chicago—looking to consumer welfare as the guiding principle of antitrust and using economic theory as the guide—is now second nature to most thoughtful commentators, judges, and practitioners of antitrust. Kodak is ironically perhaps the best proof of Chicago’s
perspective the Court's information cost theories proved merely that
the majority could identify an economic theory consistent with its
conclusions, not that the theory compelled them.115

Professor Fox analogizes Kodak to a case decided by the Court of
Justice of the European Communities in which a cash register
manufacturer was found liable to a firm it originally licensed as a
distributor;116 the firm also repaired cash registers and the
manufacturer supplied it with parts to do so.117 The manufacturer
eventually built up its own distribution network and offered a new
agreement to the distributor, which refused it because the margins
were too low.118 The manufacturer found others willing to accept the
distribution contract and, not surprisingly, thereafter refused to supply
machines or spare parts to the original firm.119 These acts, which of
course enhanced the value of the distribution agreement to the firm
that accepted it, were the basis of a suit by the original distributor. The
commission agreed that a cognizable market existed for spare parts120

success. Had it been decided 20 years ago, the Court likely would have been satisfied with the
platitudes and hasty assumptions that characterized opinions such as United States v. Topco
Associates, Inc., 405 U.S. 596 (1972), which condemned an agreement among small grocers to
develop and market private-label foods that was almost certainly procompetitive, and Fortner
Enterprises, Inc. v. United States Steel Corp., 394 U.S. 495 (1969), which allowed a claim to
proceed on the theory that the defendant tied credit for prefabricated houses to the houses
themselves. That the Kodak Court felt compelled to articulate an economic justification for its
result, cast to some extent in the form of harm to consumers rather than competitors, is a major
improvement in antitrust for which Chicago deserves substantial credit. (It would admittedly
have been better if the Court's analysis had been more realistic, but the progress is welcome
nonetheless.) But Chicago never advocated static analysis, nor does it require analysts to turn a
blind eye to market facts; no one with even passing familiarity with Coase's work would contest
the point. Though some analysts working from a Chicago orientation may have relied too heavily
on theory, that is a defect in the application of theory, not the approach itself. Nobody is perfect,
and the sins of Chicago analysts (if sins there are) pale compared to what came before. The post-
Chicago movement is occasionally both ideological and overtly anti-theoretical, which ignores
the simple reality that without a rigorous, coherent theoretical basis for determining what
antitrust goals should be and how they are likely to be achieved, it is impossible to decide cases.
One is left with lengthy recitations of facts followed by a conclusion, with no theory to distinguish
significant from insignificant facts, nor to explain why the conclusion is justifiable in terms of
antitrust policy goals. See United States v. United Shoe Machinery Corp., 110 F. Supp. 295, 299

115. Fox, supra note 105, at 760.
116. Id. at 762-66. The case was Case 22/78, Hugin Kassaregister A.B. v. Commission, 1979
117. Fox, supra note 105, at 763.
118. Id.
119. Id. The manufacturer also prohibited its subsidiaries and other authorized dealers from
dealing with the original firm as well. Id.
120. Id. at 765.
and that "a firm in a dominant position 'cannot deny its customers freedom of choice.'" From this and other cases Professor Fox extrapolates the EC rule as being that a manufacturer must supply spare parts to customary buyers so long as demand exists, unless the manufacturer can demonstrate an "objective justification" for the conduct; enhancement of consumer welfare would qualify as such a justification. The important point to Professor Fox is that a claimant need not cast its theory in terms of consumer welfare: in the case she recounts "there was no legal need to argue that customer exploitation would occur. Indeed, the facts unabashedly belied exploitation." On this point, at least, there would likely be wide agreement.

The overall argument lacks appeal, however. Officers and employees of ISOs, like all officers and employees, devote their earning power (human capital) to a single firm; they thus have relatively larger portions of their overall capital invested in the firm than do consumers and investors who own diversified portfolios. Officers and employees are therefore interested in obtaining rulings that maximize the value of their firm, even if the legal rule necessary to achieve this goal would diminish the value of firms (and investments in them) generally. This is so because their firm, and to some extent its officers and employees, will enjoy all the benefits of the wealth transfer while suffering only a very diffuse portion of the overall loss in wealth. Their interests thus conflict with the interests of consumers and investors generally, just as the interests of corporate managers in building empires (inefficient diversification at the firm level) conflict with the interests of shareholders. From this perspective the ISOs, like all competitors, cannot be considered faithful guardians of antitrust's interest in systemic efficiency. This being the case, it is

121. *Id.* at 764 (quoting [1979] E.C.R. at 1884). The Court of Justice ultimately disposed of the case on unrelated grounds. *Id.* at 765.

122. *Fox*, supra note 105, at 765.

123. *Id.*

124. *Id.* at 766.

125. Even if the officers and employees did not see the benefits directly, as in the form of raises or bonuses, they might see it in the form of a continuation of a firm that, absent the invocation of antitrust to protect it, would have expired.

126. This conclusion is not as drastic as it may seem. Professor Kauper, for example, a former head of the Antitrust Division, has recommended curtailing the ability of competitors to bring private actions under the antitrust laws. See Edward A. Snyder & Thomas E. Kauper, *Misuse of the Antitrust Laws: The Competitor Plaintiff*, 90 Mich. L. Rev. 551, 596-98 (1991). A note of caution is warranted in the context of this article, however; the empirical study on which this conclusion was based did not include tying or essential facilities claims. *Id.* at 597.
hard to understand why it would be beneficial to allow competitors to bring suit without alleging harm to overall consumer welfare and to require firms to defend each case on the merits by establishing an objective justification for their conduct. Such a rule would, as *Kodak* likely will, simply increase the number of cases in which an antitrust violation is found though consumer welfare is not harmed (the judicial equivalent of a "false positive" in statistics), while enormously increasing the cost of litigation producing such findings. If anything, one would hope the law would gravitate in the other direction, imposing what Judge Easterbrook has referred to as "filters" to attempt to weed out cases that seem unlikely as a class to enhance consumer welfare.

There is no reason to believe that antitrust has a comparative advantage in solving the problems the *Kodak* Court identified. It is important to remember that antitrust is only one method of controlling business behavior; most obviously, antitrust is an alternative to market self-correction and governmental regulation. In a proper analysis, antitrust should also be compared to other forms of legal constraint. In order to apply antitrust faithfully to enhance consumer welfare, we must ask if the benefits of applying antitrust to any given situation (including expense from unjustifiable suits brought in the hope that an unclear doctrine will be mistakenly applied, or a clear doctrine expanded beyond the reaches of its logic, and the likelihood of erroneous findings of liability) exceed the harm to the economy avoided by such application. The harm in this analysis should itself be discounted to the extent it can be ameliorated by other legal rules. Thus, if a rule other than antitrust can potentially remedy 90% of the harm identified in a particular case, an antitrust court

---

127. Of course, if one conceives of the antitrust laws as in part contemplating wealth transfers from consumers to competitors, as Professor Fox apparently does, such findings would not be false positives. It would of course be more efficient to effect such transfers through taxation or other more direct methods, but those might be politically unpalatable. On the other hand, evidence that such wealth transfers not be obtained through more open devices would seem to be an independent ground for opposing the relatively more obscure use of antitrust rulings to effectuate such transfers.


129. Id.

130. Such costs include the likelihood that a given antitrust rule may be hard to apply, and thus produce false positives (either unobjectionable conduct condemned or objectionable conduct excused) and thereby deter unobjectionable business behavior.
should be concerned with the question whether elimination of the remaining 10% warrants the invocation of the antitrust laws.131

This analysis is particularly important because many legal rules, particularly contract, are an integral part of the operation of the markets antitrust oversees. Examining the availability of alternative legal remedies is, to this extent, simply an element of an appropriate market analysis. Antitrust, with its treble damages, attorney's fees, and the prospect of broad-ranging discovery into competitors' business affairs, may operate as the equivalent of an attractive nuisance within the law. As such, antitrust may induce firms to bring suits where no objective claims exist, or to distort the operation of legal rules ingrained in the operation of markets, such as contract, by granting parties rights they could have bargained for but didn't.

Antitrust's principal competitor in *Kodak* is contract, which seems a superior way of addressing the problems the Court identified. Antitrust is neither designed to monitor the contractual relations of particular parties nor capable of such monitoring.132 The familiar metaphors of antitrust as the constitution or "magna carta" of free enterprise are apt in at least one respect: antitrust deals with the economy as a whole, and with business practices in general, not with opportunistic practices between particular firms. Yet when the infirmities in the Court's analysis of information costs are recognized, *Kodak* reduces (at worst) to a case of contractual opportunism: the only parties likely to be harmed in a systemic way by Kodak's policies were consumers who purchased Kodak machines before the policy change was announced, and thus were locked into the Kodak aftermarket by primary market switching costs.133 Such costs could allow Kodak to extract higher prices for service than the consumers would have paid had the ISOs been able to obtain parts. Though this class of consumers could theoretically have remedied the problem

131. The harm to applying other legal doctrines, comprising the same factors described in the preceding footnote, should similarly be weighed in this analysis.

132. Shapiro, *Aftermarkets*, supra note 77, at 496 ("Antitrust law is not a nimble policy instrument for fine tuning relationships between manufacturers and customers in ways that contracts cannot.").

133. Without crediting the Court's information cost analysis entirely, it is conceivable that some firms would not inquire into Kodak's service pricing prior to purchasing a copier even after Kodak's policy was announced. With respect to such firms, the policy question is whether antitrust is properly used as a consumer protection statute (in essence compelling disclosure) to protect such firms. As noted in the text, antitrust is not properly concerned with the distribution of wealth, and this principle should foreclose any use of the antitrust laws to protect firms from themselves.
through *ex ante* contracting, by negotiating for service and parts as part of the purchase of the copier, they would not have been able to learn of Kodak's policy, and thus would have been less likely to negotiate for such terms.

Kodak is in a position analogous to sailors who seek to renegotiate their wages upward when the ship is at sea and the captain cannot hire a replacement crew. Even if the captain agrees, contract will disregard the renegotiated wage term and enforce the original deal. Consumers locked into purchasing supracompetitive copier service could seek relief under analogous contractual provisions, such as the doctrine of good faith and fair dealing, in an effort to retain the benefits they thought they had obtained through their bargain. Under the relatively flexible and yet narrow principles of contract, such plaintiffs may or may not win such cases, though if they did not even have a valid contract claim the warrant for invoking antitrust would seem weaker still. At any rate, the invocation of a doctrine designed to deal with the particular contractual relations of particular parties would seem superior in every respect to the invocation of antitrust by the *Kodak* court.

With respect to the Court's information cost argument, even prospective purchasers who had no knowledge of Kodak's practices, but were simply shrewd and worried about becoming victims of opportunism in the future, would, as Professor Shapiro has rightly noted, have an array of contractual bargaining strategies available to protect themselves. These include moving the purchase of service from the aftermarket to the primary market through the purchase of a warranty with the copier, leasing instead of purchasing (thus reducing the period available for opportunism), and negotiating the right to

135. *Id.* at 97-98; See Alaska Packers' Ass'n v. Domenico, 117 F. 99 (9th Cir. 1902) (holding renegotiated wage term unenforceable for lack of new consideration).
136. Indeed, the availability of contract as a superior remedy provides the true basis of the considerable appeal of Justice Scalia's dissenting opinion in *Kodak*, 504 U.S. at 486-502. Justice Scalia began from the premise that if Kodak had bundled the copier, parts, and service *ex ante*, there would have been no question that its conduct was legal. *Id.* at 490-91. This is true, but not, as Justice Scalia argues, because there could have been no tie. If one accepts that copiers, parts, and service are distinct antitrust markets, a tie is theoretically possible. Such bundling would, however, unquestionably have cured any defects in consumers' information regarding the lifecycle pricing of copiers—they would have priced for all three markets because buyers would have been in all three markets simultaneously. This pricing would have taken place at the time of the initial copier purchase, which occurred in the concededly competitive primary market. On this point, see Lee v. Life Insurance Co. of North America, 23 F.3d 14, 20 (1st. Cir. 1994) (noting curative effect of *ex ante* disclosure).
receive parts and service on the same terms offered to new copier purchasers (i.e., most favored nation clauses). Admittedly, not all purchasers are shrewd, but antitrust can hardly remedy that problem. And there must be some limits to the presumed shortcomings of the reasoning power and negotiating acumen of consumers. No legal doctrine short of explicit price regulation could protect consumers from themselves. In particular, there is absolutely no reason to believe that antitrust can do better than contract.

Moreover, there is a serious reason to question whether application of antitrust principles to these cases could solve the problems antitrust ought to solve. Even if the ISOs remained players in Kodak's aftermarkets, it does not follow that pricing in the aftermarkets will be competitive. If Kodak uses its inevitably strong position in its aftermarket to charge supracompetitive prices, which is the ISOs' entire theory, it may well provide a price umbrella under which the ISOs could price below Kodak but still at a supracompetitive level. The ISOs' interests are served by charging the highest price they can, not necessarily in pricing at marginal cost. In markets with few ISOs, the ISOs would have no reason to undercut Kodak's price more than a little. Is the difference between a 15% overcharge and a 20% overcharge in the aftermarkets, when properly accounted as only a portion of the total copier cost to the consumers, worth invoking the antitrust laws, particularly when contract could provide a substantial if not complete remedy to the problems the Court identified?

137. See Shapiro, Aftermarkets, supra note 77, at 488-89.
138. And price regulation may simply substitute a political market for an economic market.
139. I have often heard the argument made that, whatever consumers should do in theory, they simply do not price copiers such as Kodak's on a lifecycle basis, and therefore antitrust must step in. Accepting the factual predicate as true for purposes of analysis, the conclusion that antitrust should be invoked does not even begin to follow. There are good reasons why legal penalties and incentives should be structured on an assumption of rational behavior. Among other things, such structures make rational behavior a superior strategy to irrational behavior. More pragmatically, consumers may often (if not almost always) fail to read in full the insurance policies, automobile purchase agreements, and other contracts they sign. That does not mean that a failure to read an agreement should be a defense to an action for breach of contract.

The legal supposition that consumers have agreed to all the terms of the contracts they sign, and thus may be sued for breach of such terms, provides incentives to specify all the relevant rights and obligations ex ante, thus reducing the uncertainty (and thus the risk) inherent in contracting. If consumers could avoid their contracts simply by avowing that they had not read them, the risks of contracting, and thus the costs of contracting, would increase, resulting in a reduction in welfare from foregone opportunities for exchange. The result in Kodak has much the same effect.

140. Kodak, 504 U.S. at 465.
On balance, the answer is no: as a matter of antitrust policy, *Kodak* was wrongly decided. Though the case identified a relatively narrow group of consumers that suffered actual harm from the practices at issue, and thus has a legitimate relationship to consumer welfare, the harm was not easily amenable to cure through application of antitrust. It is at least probable, if not certain, that the costs of the decision, in the form of meritless litigation expenses, economically unjustifiable settlements and verdicts, and legitimate business practices wrongly foregone on the advice of counsel quite reasonably unable to distinguish between safe and unsafe practices, will exceed the relatively narrow benefits.\(^1\)

C. The Proper Scope of the Essential Facilities Doctrine

Where does all this leave the essential facilities doctrine? *Kodak*’s ruling that aftermarkets may constitute relevant antitrust markets\(^2\) has the effect of multiplying the number of cases in which a firm could be said to have an actionable level of market power, regardless of competition in the primary market. If, therefore, power in markets such as “parts or service for Kodak copiers” is sufficient to state an essential facilities claim, the antitrust laws could be transformed into a powerful compulsory dealing law. Fortunately, neither analysis nor precedent supports such a result. *Terminal Railroad, Otter Tail*, and the prominent circuit cases such as *MCI Communications Corp. v. AT&T* and *Olympia Equipment Leasing Co. v. Western Union Telegraph Co.*, limit the scope of the doctrine to cases in which the defendant possesses a natural monopoly. This includes only cases in which the relevant market could not support more than one firm and thus one in which the facility could not be replicated.\(^3\)

Limitation of the doctrine to situations involving natural monopoly is not only consistent with the important antitrust principle, confirmed in *Aspen Skiing*, that a firm has no obligation to assist its

---

141. The decision has thus far cost Kodak $72 million. It is worth noting in this regard that Kodak has recently announced plans to sell its copier division, citing stiff competition and low profits. See *Kodak Planning to Sell Copier Division: Poor earnings, Stiff Competition Cited*, WALL ST. J., Jan. 16, 1996, at B9.


143. See supra text accompanying note 32 (defining natural monopoly); Fishman v. Estate of Wirtz, 807 F.2d 520, 574 (7th Cir. 1986) (Easterbrook, J., dissenting) (“All of the essential facilities cases involve natural monopolies.”). The doctrine could conceivably be applied in cases where a firm had a monopoly by reason of some governmental regulation barring entry into a relevant market, which was not a natural monopoly, but such cases would have to be addressed under whatever regulatory scheme erected the barrier in the first instance.
competitors, it is the only way to achieve doctrinal coherence without vitiating that principle. By hypothesis, in a natural monopoly market it will not be profitable for competitors to replicate existing systems, such as the St. Louis terminal facilities or Otter Tail’s power lines. Thus, “[i]f the antitrust laws allowed the owner of the facility to turn users away, there would be an incentive to wasteful duplication. Monopoly profits serve as a lure; the result would be inefficient.” It is for this reason, which is directly related to consumer welfare and ineradicably tied to the cost structure of the market, that access may be properly compelled in certain cases. At least in “bottlenecking” cases such as Otter Tail, such a rule could increase competition and thus efficiency.

In situations not involving natural monopoly the market will support more than one firm. Once this determination is made, the essential facilities doctrine has no role to play because an entrant may replicate the facility in question. So long as replication is possible, the claim that a competing facility would be too costly to build is entitled to no weight. Such arguments represent a shift from the claim that it is economically infeasible for a firm to replicate a facility (such as a

145. Fishman, 807 F.2d at 574. This statement needs some qualification; obviously if firms knew ex ante that a market would only support one firm any incentives to duplication would likely be outweighed by the certainty of failure. Nevertheless, there would be at least some risk from firms that thought they could succeed against the odds. Efficiency would dictate that access would be a better solution in such situations.
146. This would only be true, however, if the firm in possession of the monopoly received a fair rate of return: possession of a natural monopoly is no warrant for compulsory subsidization of competitors at the expense of investors. This illustrates the importance of the federal regulatory scheme, with its rate of return calculations, to the result in Otter Tail.
147. Twin Labs., Inc. v. Weider Health & Fitness, Inc., 900 F.2d 566, 568 (2d Cir. 1990) (“Antitrust law, however, does not require one competitor to give another a break just because failing to do so offends traditional notions of fair play . . . . At the very least, a plaintiff must demonstrate that 'duplication of the facility would be economically infeasible.'”) (quoting Hect v. Pro-Football, Inc., 570 F.2d 982, 992 (D.C. Cir. 1977), cert. denied, 436 U.S. 956 (1978)); Fishman, 807 F.2d at 574 (Easterbrook, J., dissenting) (“The high cost of a facility does not make it 'essential.'”); HOVENKAMP, ANTITRUST POLICY, supra note 9, at 275 (“The courts have generally interpreted the essential facilities doctrine to require a showing that no practical alternatives are available, including alternatives that face cost disadvantages.”). Assuming efficient capital markets, plaintiffs seeking to undertake positive net present value investments involving the replication of an existing facility, even at great cost, should be able to finance such investments, though the cost might be high. The cost of financing is not a legitimate antitrust concern, however (among other things it is a distributional concern). And if an investment involving replication did not have a positive net present value, and thus could not sustain financing, it is hard to see why social welfare would be enhanced by allowing the investment to go forward through a competitor subsidy.
bridge or power grid) to the claim that a plaintiff firm would be better off, and thus a better competitor, if it did not have to bear the capital expenses borne by the defendant. This immediately reduces to the somewhat less than lofty legal insight that competition is easier if one is entitled to subsidies from one’s competitors. Such claims are entitled to no legal weight, either theoretically or as a doctrinal matter. In appropriate cases (which will be very few), the cost of replication may have a role to play in analyzing the likelihood of entry when considering whether an existing firm has market power. Such analysis is fundamentally different than an “essential facilities” analysis, however, and will not support an order that a firm deal with its competitors. By reducing the returns available to existing firms, compulsory dealing would diminish investment ex ante in otherwise profitable ventures and thus diminish consumer welfare.

As we shall see, this analysis compels the rejection of essential facilities claims in most contexts relevant to computer software. Before explaining this conclusion, however, we must examine some of the peculiar characteristics of copyright as applied to software.

III

Networks, Externalities, and Innovation: Competitive Implications of Copyright Protection of Computer Software

Having defined the proper parameters of the essential facilities doctrine, we turn to the application of that doctrine in the rapidly-evolving field of copyright protection of computer software. As noted generally above, copyright’s concerns parallel those of antitrust in some respects with several important differences. To recap briefly, the production of intellectual property, which is the statutory and constitutional grail of copyright, is largely a function of the scope of protection against copying provided by the copyright laws. If the protection is too strong (or too broad, if you prefer) investment in

148. See Olympia Equip. Leasing Co. v. Western Union Tel. Co., 797 F.2d 370, 379 (7th Cir. 1986) (“Consumers would be worse off if a firm with monopoly power had a duty to extend positive assistance to new entrants, or, having extended it voluntarily a duty to continue it indefinitely.”); see also Reisner v. Gen. Motors Corp., 671 F.2d 91 (2d Cir.), cert. denied, 459 U.S. 858 (1982).

149. Professor Areeda makes a similar point. See Areeda, supra note 19, at 851.
intellectual property will be diminished by the risk of infringement, including the threat of infringement suits. If the protection is too weak (narrow), investment will be diminished by the risk of competition from copiers who did not bear the costs of innovation.

Adjusting the scope of legal protection necessary to achieve the optimal level of production and innovation (investment) in intellectual property is the key to a successful copyright regime. This balancing may call for "fine-tuning" the property rights granted by copyright, at least in a rough sense. Cookbooks may receive less protection than novels, for example, because less investment is required to publish them and the lower returns from relatively weak copyright protection thus may not decrease production in cookbooks in any material way.\textsuperscript{150} Antitrust has no analogue to this endeavor. It takes the allocation of property rights—whether it be the right to exclude others from the plant premises or to exclude others from copying software—as a given and inquires whether firms possessing those rights are engaging in anticompetitive activity, by which antitrust means activity that will allow a firm or firms to reduce output and price materially above marginal cost. The analytical tools necessary to this analysis, gathered under the general heading of price theory, provide few if any clues to the question of how strong a copyright should be.\textsuperscript{151}

We begin our examination of software with some simplified rudiments of technology.\textsuperscript{152} Computer hardware refers to wires, circuits, chips, and the beige boxes that contain them.\textsuperscript{153} Software refers to written code that instructs the computer to perform various functions.\textsuperscript{154} We will be concerned with two types of software:

\textsuperscript{150} See Feist Publications, Inc. v. Rural Tel. Serv. Co., Inc., 499 U.S. 340, 349 (1991) ("the copyright in factual compilations is thin"); Julie E. Cohen, Reverse Engineering And The Rise of Electronic Vigilantism: Intellectual Property Implications of 'Lock-Out' Programs, 68 S. CAL. L. REV. 1091, 1108 (1995) (arguing that copyright for software should be analogous to factual compilations such as cookbooks). This is not to say that the calculation of returns is an exact science, which it of course is not. Firms, which are far better at estimating the returns necessary to undertake a project in intellectual property, make mistakes all the time. And all investments have some degree of risk, which is to say that more things can happen than will happen; an investment may thus turn out bad even if it has a high expected value.

\textsuperscript{151} See HOVENKAMP, ANTITRUST POLICY, supra note 9.

\textsuperscript{152} The following discussion draws upon Professor Menell's review of technology. See Peter S. Menell, Tailoring Legal Protection For Computer Software, 39 STAN. L. REV. 1329 (1987) [hereinafter Menell, Tailoring]; Peter S. Menell, An Analysis of the Scope of Copyright Protection for Application Programs, 41 STAN. L. REV. 1045 (1989) [hereinafter Menell, Scope of Protection].

\textsuperscript{153} Menell, Tailoring, supra note 152, at 1333.

\textsuperscript{154} Principal types of code include object code, which refers to a series of binary units (0's and 1's) that can be read by a computer directly and that instructs electronic switches to be
operating systems and application programs. Operating systems software refers to code that manages the internal functions of the computer, such as retrieving data from or storing it in the computer's memory, communicating with printers and modems, and the like.\textsuperscript{155} Application software refers to code that performs specific functions, such as word processing, spreadsheet analysis, and bookkeeping.\textsuperscript{156} The scope of protection given to computer software under the copyright laws is the key to understanding the issues raised by our first two hypotheticals. \textit{MAI Systems Corp. v. Peak Computer, Inc.}\textsuperscript{157} and \textit{Data General Corp. v. Grumman Systems Support Corp.}\textsuperscript{158} typify the problems posed by the first hypothetical case.

A. \textbf{Use of Copyright As A Barrier To Competition: ISO Copying Of Manufacturer Software}

Data General manufactured and serviced mini-computers;\textsuperscript{159} Grumman Systems Support Corporation was an ISO that competed with Data General in the market for service of computers manufactured by Data General,\textsuperscript{160} which had the lion's share of the market for servicing computers it had built.\textsuperscript{161} Data General's dispute with Grumman revolved around a diagnostic software program called Adex, which worked with Data General's most advanced and profitable (both in sales and service) line of computers.\textsuperscript{162} Data General would license Adex to customers who serviced their own computers, but would not license Adex to customers who purchased service, nor to ISOs\textsuperscript{163} such as Grumman.\textsuperscript{164} Data General refused to

\begin{itemize}
  \item[155.] \textit{Id.}
  \item[156.] \textit{Id.}
  \item[157.] 991 F.2d 511 (9th Cir. 1993), \textit{cert. dismissed}, 510 U.S. 1033 (1994).
  \item[158.] 36 F.3d 1147 (1st Cir. 1994).
  \item[159.] \textit{Id.} at 1152.
  \item[160.] \textit{Id.}
  \item[161.] "Although DG [Data General] ha[d] no more than a 5\% share of the highly competitive 'primary market' for mini-computers, DG occupie[d] approximately 90\% of the 'aftermarket' for service of DG computers." \textit{Id.}
  \item[162.] \textit{Id.} at 1154.
  \item[163.] \textit{Id.} at 1152. In \textit{Data General}, the term used is TPM, "third party maintainer," which is the equivalent of ISO, "independent service organization," in \textit{Kodak}, 504 U.S. at 451.
  \item[164.] \textit{Data General}, 36 F.3d at 1152.
\end{itemize}
license Adex to Grumman, placing Grumman at a competitive disadvantage and prompting what might charitably be called a creative campaign by Grumman to acquire copies of the software. Data General responded by suing Grumman for violation of its copyrights and misappropriation of trade secrets; Grumman defended in part by contending that Data General was illegally maintaining a monopoly in the market for service of its computers by refusing to license Grumman to use the diagnostic software.

MAI built computers and designed software to run them, including both operating systems software and diagnostic software. MAI licensed this software to purchasers of its machines; the licenses allowed purchasers to make a specified number of copies of the software for their own internal needs, but prohibited third parties from making any copies. Peak was an ISO that serviced MAI computers. It was founded by former MAI service employees, who took some of their customers with them when they started Peak. It was undisputed that to service the computers Peak needed to start them up and determine whether they were running the MAI operating system properly. This entailed copying the operating system software from a storage device, such as a hard disk, to the computer's random access memory (RAM), which is comprised of memory chips that retain the information they are given only so long as they are electrically charged. The Peak technicians could then view the

---

165. Id.
166. Grumman's methods created a poor record, to say the least. Former Data General employees took copies with them when they moved to Grumman; Grumman used copies given to former Data General service customers who terminated Data General's service but refused to return the software (as required by Data General's service agreement); and Grumman used and copied software inadvertently left behind by Data General technicians. Id. at 1154-55.
167. In due course a jury awarded Data General over $27,000,000 in damages. Id. at 1152.
168. Id. at 1166.
169. MAI, 991 F.2d at 516.
170. Id. at 517.
171. Id.
172. Id. at 513.
173. Id.
174. Id.
175. As a district court in another case involving MAI put it:
[T]he RAM in MAI computers is composed of dynamic random access memory chips (DRAMs) and static random access memory chips (SRAMs), with DRAMs predominating. DRAMs are 'dynamic' because they must be repeatedly 'refreshed' with an electrical charge or they will lose the information stored within them. SRAMs do not require such refreshing. Both DRAMs and SRAMs, however, quickly lose their information when electrical power is removed. Thus, they are volatile, and for this
"systems error log," which was a part of the operating systems software, to diagnose the fault in the system and subsequently verify the effectiveness of a repair.\textsuperscript{176}

MAI sued Peak, claiming that this copying violated MAI's copyrights, and obtained an injunction that was affirmed by the Ninth Circuit.\textsuperscript{177} The Court's analysis was straightforward: the Copyright Act prohibits copying, which is defined as fixing a work in a tangible medium in such a manner that the work can be "perceived, reproduced, or otherwise communicated, either directly or with the aid of a machine or device."\textsuperscript{178} Peak's copying of the operating system into the computer RAM, for the express purpose of "perceiving" and acting on the information generated, easily met this definition.\textsuperscript{179} Given the pleadings before it, the court did not believe it needed more information to decide the case. This was unfortunate, because the decision thus did not reach the more interesting issue of whether Peak could defend its infringement on the grounds of fair use. If Peak tried, could MAI have rebutted the defense on the ground that Peak was formed by former MAI employees, arguably free riding on MAI's investment in computer technology and in them? Or would the court have rejected such an argument on the theory that if MAI had wanted to bar its employees from competing, it should have negotiated enforceable noncompetition agreements with them \textit{ex ante}? What about the consumers—purchasers of computers who presumably knew the terms of the license when they bought the computers?\textsuperscript{180} Should the law leave the fate of the ISOs to the \textit{ex ante} contracting skills of purchasers, or should it intervene on their behalf?

From the antitrust perspective, these ISO cases are familiar; they bear a strong resemblance to \textit{Kodak}, and indeed rely almost entirely on the Court's approval in that case of antitrust suits in aftermarkets. There is a crucial difference. Whereas Kodak excluded ISOs from its aftermarkets through the alleged use of pressure on parts suppliers to

\begin{itemize}
\item reason are often termed 'volatile memory,' in contrast with ROM, the contents of which are semi-permanent.
\item \textit{MAI}, 991 F.2d at 518.
\item \textit{Id.} at 513.
\item \textit{Id.} at 517 (quoting 17 U.S.C. § 101 (1990)).
\item \textit{Id.} at 518.
\item In this respect, \textit{MAI} presents a harder case than \textit{Kodak} for invocation of the antitrust laws because in \textit{MAI} the license was specified at the time of sale. Thus the case is analogous to Justice Scalia's hypothetical contract for a Kodak copier in which it is specified \textit{ex ante} that only Kodak may service the machine. \textit{Kodak}, 504 U.S. at 490-91.
\end{itemize}
tie parts to service, MAI and Data General sought to exclude ISOs in their cases through suits alleging copyright infringement. The antitrust claims in these cases thus tend to raise essential facilities claims rather than tying claims, based on the notion that access to the diagnostic software is essential for the ISOs to compete in service markets.

This difference is important in two respects. First, whatever one thinks of the merits of the tying theories approved in *Kodak, MAI* and *Data General* lack the elements of legitimate essential facilities claims. Second, apart from this substantive failure, antitrust should defer to such "power" as copyright laws confer upon copyright holders even if, as in these cases, the policy goals of copyright are not furthered.

With respect to the elements of the essential facilities claim, the threshold question is whether the market for the essential facility—here the diagnostic software to which the ISOs sought access—bears the characteristics of a natural monopoly such that replication of the facility is not economically feasible. Stating the question in these terms reveals the odd fit between the essential facilities doctrine and the *Kodak* aftermarkets in which the ISOs operate. In the first place, which market is diagnostic software a part of? For MAI's computers, the "error log" portion of the diagnostic software was part of the computer's general operating system software; other portions of MAI's diagnostic software, like Data General's Adex software, were separate. At least the built-in operating system software would appear to be part of the computer itself for market analysis purposes, and thus part of the competitive primary market. It is a fair question whether even Adex, which was sometimes licensed apart from Data General computers, can constitute a meaningful market. Though the district court found a factual dispute on the point precluding summary judgment, it is not at all clear that it would be efficient for Data General to produce diagnostic software separately from the computers themselves.

---

181. See *supra* note 147 and accompanying text.
182. Apologies to those who concur with Churchill that a preposition is a terrible thing to end a sentence with.
183. *MAI*, 991 F.2d at 518.
184. *Id.* at 517.
185. The court concluded that diagnostic software could be a separate market from computer service, because some customers were willing to purchase the software without purchasing service from either Data General or a third party. *Data General*, 36 F.3d at 1180. This fact does not address the question whether it would have been efficient for Data General to write diagnostic software except as part of its manufacture of the computers. This is not to say that there could never be a separate market for diagnostic software. There may be such a market, but
The difference is important because the primary market for minicomputers was "highly competitive." Data General, for example, had only 5% of the market. The cost structure of the primary market thus quite obviously did not resemble a natural monopoly market in any respect. Because this threshold condition for an essential facilities claim could not be met, the claim should rightly fail. This answer, satisfactory as a doctrinal matter, and vital if the essential facilities doctrine is to be confined to economically sensible turf, still has a bit of an unsatisfactory ring. There is no realistic chance that any ISO will go into the computer market. Apart from their lack of capital and expertise at building computers rather than servicing them, computers were not an input for the ISOs' business—they were what the ISOs wanted to service.

Attempting to move diagnostic software out of the primary market does not solve the problem, either. Recall that under Kodak and Jefferson Parish, a market for a good or service exists if there is sufficient consumer demand for firms to provide that good or service apart from related goods or services. Thus, tires are a separate market from cars because there is sufficient demand to support tire manufacturers on their own, even though there is an obvious complementary relationship between the two markets. The same is true for cameras and film, VCRs and videocassetes, CD players and CDs, etc. But the cases reveal no evidence of a market for minicomputer diagnostic software. MAI and Data General only produce such software in conjunction with (and largely as a part of) their computers; the ISOs in those cases did not produce such software at all.

Even assuming that diagnostic software could be separated from computers in an economically meaningful way, it would not necessarily follow that the ISOs could not write diagnostic software

---

186. Data General, 36 F.3d at 1152.
187. Id.
188. One could of course imagine an ISO asserting a perverse reverse tying claim—that MAI had tied diagnostic software to operating systems software in order to move diagnostic software into a competitive market—but such a claim would have no doctrinal support. The mind does reel, however, at the very odd arguments that begin to make sense with respect to these unusual claims.
themselves or that an adept programmer would not see an opportunity to write programs for different brands of computers for sale to ISOs. At least one court has felt comfortable asserting that ISOs "are free to develop their own diagnostic software, should they so choose."191 This conclusion may be correct, but we must be more careful than the court was in drawing it. Software has high fixed costs (such as employing programmers to write the code) and very low variable costs (such as copying the code onto disks and selling them). Assuming that different software is required for each brand of computer, the market over which the fixed costs can be spread may be fairly small (though perhaps not for third-party suppliers), thus making it difficult for a party that did not reap revenues from the computers themselves to justify incurring the costs of writing software.192 While this would tend to support an argument that the cost structure of diagnostic software made it difficult to replicate, it also tends to undermine the claim that diagnostic software is a facility separate from computers in any meaningful sense.

There is thus little merit to the claim that access to diagnostic software in a computer aftermarket is an essential facility under the Sherman Act. Even if diagnostic software could realistically be considered a facility independent of the computers, with natural monopoly characteristics, considerations of copyright policy compel the rejection of the antitrust claim. As the Data General court rightly noted, copyright grants a "monopoly" over copying of the relevant work (which is of course not to be confused with a monopoly in a relevant market for antitrust purposes) in order to ensure that sufficient economic returns are held out to induce firms to invest in creative works.193 Antitrust is only concerned with returns to particular firms to the extent that it seeks to prevent firms from pricing above their marginal cost because of some defect in market structure, not because of the characteristics of the firm's assets.194 Because copyright is concerned with securing adequate returns to creators of creative works, the "power" granted to copyright holders must be determined

191. MAI, 845 F. Supp. at 368.

192. On the other hand, it could be that computer firms would sell hardware at cost or at a slight loss, planning on earning a profit solely through service, protected by copyright. The possibility of revenue shifting between primary and aftermarkets to take advantage of unique aftermarket elements is examined in greater detail in the discussion below concerning network effects.

193. See generally Data General, 36 F.3d at 1187.

194. See DOJ INTELLECTUAL PROPERTY GUIDELINES, supra note 3, at § 2.2.
in accordance with copyright principles and copyright goals, through application or amendment of the copyright laws, not through the courts, and particularly not through courts using the antitrust laws. After all, the ISOs in these cases have problems because copyright prohibits them from copying the diagnostic software temporarily into the computers’ RAM, not because of any structural defect in the aftermarkets, as was alleged in Kodak. Moreover, antitrust has traditionally deferred to the intellectual property laws with respect to the breadth of intellectual property rights. Such deference is appropriate, even in these cases, because antitrust does not have the analytical tools necessary to determine the appropriate breadth of such rights.

Still, there are very good arguments, based in the policy and goals of copyright, for amending those laws to allow the limited, temporary copying of diagnostic software into computer RAM at issue in these cases. Admittedly, both MAI and Data General involved the copying of material portions of copyrighted software in order to allow the copier to use the expressive aspects of that software to compete with the copyright holder. Neither Peak nor Grumman was attempting to create new software or other innovations; they simply wanted to use copyrighted work as an aspect of their provision of copier service. The courts’ condemnation of their copying thus


196. As of this writing, a bill (H.R. 533) has been introduced to remedy the problem discussed in the text. The bill would modify Section 117 of the Copyright Act to permit a “rightful possessor” of a program to make or authorize the making of a copy of the program so long as the copy is “created as an essential step in the utilization of the computer program in conjunction with a machine . . . .” (Other modifications of Section 117 have also been proposed.)

197. Given the nature of RAM, of course, the copy would disappear when the RAM chips were no longer charged with electricity—i.e., when the computer was turned off.

198. It was on this basis that the court in Triad Systems Corp. v. Southeastern Express Co., 64 F.3d 1330, 1337 (9th Cir. 1995), cert. denied, 116 S. Ct. 1015 (1996), rightly distinguished these cases from the its holding in Sega Enterprises, Ltd. v. Accolade, Inc., 977 F.2d 1510, 1514 (9th Cir. 1992). The copying in Sega was for the purpose of allowing the defendant to enable its own creative work, which did not infringe Sega’s copyright, to interface with Sega’s hardware (the elements of the software code necessary to achieve compatibility were not themselves copyrightable). Id. In contrast, neither Peak nor Grumman intended to create any intellectual
resulted in little (and likely no) loss of creative work to the public. To this extent, one may say that the decisions are consistent with the goals of copyright: if an alleged infringer's copying does not promote the creation of new intellectual property, why should such copying be allowed?

One reason is that this logic cuts both ways. With respect to the software at issue, MAI and Data General had no more interest in expanding the scope of available intellectual property than Peak or Grumman. They simply wanted to be able to fix the computers, too. Assertion of their copyrights thus fits awkwardly with the purposes of copyright, even though their programs were indisputably copyrightable (and copyrighted) material. By allowing the use of copyrighted material as a utilitarian tool for the operation of nonexpressive hardware, copyright does not increase "the harvest of knowledge," but may give the copyright holder the ability to extract greater returns from its hardware through the service market over the life of the hardware. While copyright is concerned with preserving those returns related to the creation of information, there is little reason to allow the use of copyright to secure supracompetitive returns unrelated to that purpose.

As in Kodak, the costs of the strategic use of copyright as a barrier to entry into the service aftermarket for a computer falls in large part on the purchaser of the computer. Until Congress acts in this regard, consumers who own machines that entail aftermarkets in which copyrighted material plays a part may attempt to protect themselves from price gouging through the same panoply of measures available to purchasers of Kodak's copiers. In healthy markets, the adverse reputational effects of using copyright to extract an extra pound of flesh from a purchaser of a durable good is a significant deterrent to such behavior. Assuming a competitive primary market, a firm that engages in such behavior will lose sales from prospective purchasers and will have trouble persuading existing owners to buy again when the time comes. It is no coincidence, as Professor Shapiro

---

200. I say "ability" rather than "power" because the latter term has the seemingly hypnotic ability to sidetrack meaningful analysis and create the illusion of problems properly remediable under the antitrust laws. As the previous discussion should establish, such "power" as exists in aftermarkets is likely to harm consumers only in relatively rare cases, and then only on an essentially "one-shot" basis.
notes, that MAI emerges from the "declining minicomputer market."\textsuperscript{201} Firms exiting a market obviously will not be repeat players in that market, and thus will not bear the costs of opportunistic behavior that gouges the existing customer base in that market. But the failure is not complete. If firms exiting a declining market continue to compete in even loosely related markets, they will have some concern for their reputation; customers victimized by a firm in the minicomputer market will be wary of buying other items such as desktops or laptops from that firm as well. To the extent reputational forces are not a complete answer, however, change must come from within copyright itself. Antitrust lacks the analytical tools to distinguish between uses of a copyright that further the purposes of the copyright laws and uses that do not.

B. Copying For Compatibility: Reverse Engineering, Replication, And The Essential Facilities Doctrine

As essential facilities cases falling within the antitrust/copyright intersection go, MAI and Data General are relatively easy. Neither antitrust doctrine nor antitrust policy supports an essential facilities claim. Copyright doctrine supports an infringement claim, but copyright policy fairly clearly calls for modification of this rule. Thus, under current doctrine, copyright defendants who are antitrust counterclaimants should lose the infringement action and their counterclaim: on a normative level they should win the infringement action. Copyright policy and doctrine come much closer together in Sega Enterprises, Ltd. v. Accolade, Inc.,\textsuperscript{202} which poses both the aftermarket problem (there is, it seems, no end to Kodak's reach) and the significantly different copyright issue of copying software for the sole purpose of adapting noncopyrightable functional code to independent expressive work in order to render that work compatible with another system component.\textsuperscript{204} Much has been written about Sega, and the full effects of the decision are likely yet to be realized. Like Kodak, it appears to be a trend-setter. Unlike Kodak, it appears to be pointed in the right direction, and there are many good reasons to be grateful for the decision from a policy perspective.

\textsuperscript{201} Shapiro, Aftermarkets, supra note 77, at 490.
\textsuperscript{202} 977 F.2d 1510 (9th Cir. 1992).
\textsuperscript{203} Id. at 1514.
\textsuperscript{204} Id. at 1522.
The news is not all good, however. Although the result in Sega was desirable, the opinion cannot be completely squared with the Supreme Court precedent the opinion cited. Understandably, for a circuit court confronting a novel issue and wanting to reach a desirable result, the court did not acknowledge this fact nor did it attempt to justify its departure on the economic merits of the case. The court instead attempted to portray the decision as fitting within the existing law. It did not, at least not exactly. The court’s precedential constraints led it to avoid issues that must be confronted in the debate over how the fair use defense will evolve to accommodate the particular needs of computer software. Sega, in short, needs to be dissected, analyzed in the context of significant empirical research, and then embodied in the Copyright Act.

1. The Sega Decision And The Fair Use Defense

Sega manufactured computer-game hardware and computer games, which, as noted above, are simply a type of software program. In order to run on the Sega hardware, computer games needed certain codes copyrighted by Sega, which normally licensed them to computer game developers. Accolade manufactured computer games and wanted to produce games that could run on Sega hardware. Accolade explored the possibility of obtaining a license from Sega, but found the terms unacceptable. Accolade therefore began to “reverse engineer” Sega’s games in an effort to isolate the code needed to make its games compatible with Sega’s hardware. This involved decompiling the program from object code into source code, which could then be read by Accolade’s programmers. Accolade programmers studied three different Sega games, isolated similarities, and eventually isolated the code necessary for compatibility, which they then added to a manual describing the interface requirements. These portions of the code were themselves functional, and thus not entitled to
independent copyright protection.\textsuperscript{209} The program as a whole, however, all of which was copied, was entitled to copyright protection.\textsuperscript{210} Accolade contended that, apart from the interface specifications, "none of the code in its own games is derived in any way from its examination of Sega's code."\textsuperscript{211}

On these facts, Accolade was found to have copied Sega's code within the meaning of the copyright statute,\textsuperscript{212} a decision that was doubtless correct, but also to have a fair use defense to Sega's infringement claim, which is the element of the decision of interest to us here.\textsuperscript{213} The nonexclusive statutory elements of a defense of fair use 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.\textsuperscript{214}

On the first element, Sega argued that the purpose of Accolade's copying was to make a profit by competing against Sega's products.\textsuperscript{215} At the time \textit{Sega} was decided, the Supreme Court's decision in \textit{Harper & Row, Publisher's, Inc. v. Nation Enterprises}\textsuperscript{216} held that a commercial purpose weighed heavily against a finding of fair use.\textsuperscript{217} In light of this authority, the \textit{Sega} court's analysis was marvelously innovative and prescient. As the court reasoned:

\begin{quote}
[T]here is no evidence in the record that Accolade sought to avoid performing its own creative work. Indeed, most of the games that Accolade released for use with the Genesis console were originally developed for other hardware systems. Moreover, with respect to the interface procedures for the Genesis console, Accolade did not seek to avoid paying a customarily charged fee for use of those procedures, nor did it simply copy Sega's code; rather, it wrote its own procedures based on what it had learned through disassembly . . . . [A]lthough Accolade's ultimate purpose was the release of Genesis-compatible games for sale, its direct purpose in
\end{quote}

\textsuperscript{209} 17 U.S.C. § 102(b) (1995); 977 F.2d at 1522-23.
\textsuperscript{210} \textit{Sega}, 977 F.2d at 1518 (referencing 17 U.S.C. § 106(1)-(2) (1990)).
\textsuperscript{211} 977 F.2d at 1515.
\textsuperscript{212} \textit{Id.} at 1527.
\textsuperscript{213} \textit{Id.} at 1527-28.
\textsuperscript{214} \textit{Id.} at 1521-22 (quoting 17 U.S.C. § 107 (1990)).
\textsuperscript{215} 977 F.2d at 1522.
\textsuperscript{216} 471 U.S. 539 (1985).
\textsuperscript{217} \textit{Id.} at 562.
copying Sega's code, and thus its direct use of the copyrighted material, was simply to study the functional requirements for Genesis compatibility so that it could modify existing games and make them usable with the Genesis console.218

These facts, which sharply distinguish Sega from MAI and Data General, led the court to the further conclusion that Accolade's copying was consistent with the overarching purposes of copyright—to "promote the Progress of Science and useful Arts"219—because Accolade's copying "led to an increase in the number of independently designed video game programs offered for use with the Genesis console."220 This analysis decided the first element in Accolade's favor,221 and has been substantially vindicated, at least in the context of parody, though perhaps with broader implications, by the Supreme Court's opinion in Campbell v. Acuff-Rose Music, Inc.222

The court's discussion of the fourth element is more problematic. The question was whether, if the challenged practice became widespread, "it would adversely effect the potential market for the copyrighted work."223 Because copyright analysis focuses on returns to those who invest in intellectual property, and because such returns are a function of market forces, this factor is "undoubtedly the single most important element of fair use."224 The Sega court interpreted this question to ask only whether the challenged practice would "usurp the market"225 for the copyrighted work, in which case it would not be fair use, or merely facilitate competition with the copyrighted work by a defendant's product, in which case a fair use defense would succeed.226 This dichotomy is problematic, and the court's brief discussion did little to point the way towards a solution. Two principal issues are raised: how is the "market" in which the infringer's use is evaluated to be defined, and what is the "use" relevant to the fourth element.

One would hope that an analysis of market effects would begin with a reasoned discussion of market definition. That hardly ever happens, though, and in this shortcoming Sega is regrettably typical.

218. Sega, 977 F.2d at 1522.
220. Sega, 977 F.2d at 1523.
221. Id.
226. Sega, 977 F.2d at 1523.
The *Sega* court did not formally define the market in which it evaluated the fourth factor. Indeed, the court appeared not to have perceived a definitional issue. This may be understandable because the Copyright Act provides that the market relevant to a fair-use defense is the market for the copyrighted work.\(^{227}\) In many cases only one market will fit this bill; the question in *Sega* was not so easy—was the relevant market the market for computer games, in which Sega's copyrighted material competed with Nintendo's (and later Sony's) material,\(^{228}\) or was it the aftermarket for Sega-compatible computer games? The court's opinion makes clear that it perceived the relevant market to be the one for Sega-compatible games. Thus the court's reference to Accolade as a “new competitor, the first lawful one that is not a Sega licensee,”\(^{229}\) and its more worrisome statement that “an attempt [by Sega] to monopolize the market by making it impossible for others to compete runs counter to the statutory purpose of promoting creative expression and cannot constitute a strong equitable basis for resisting the invocation of the fair use doctrine.”\(^{230}\)

Professor Miller picks up on the court's “monopolization” rhetoric and takes exception with this definition, contending that:

> There is no principled basis for the Ninth Circuit's definition of the relevant market in which monopoly control was to be prevented as the market for Sega cartridges; for example, the court also could have defined the market in terms of home entertainment systems, a market in which Sega and Nintendo compete fiercely, and required Accolade to design and vend its own console, or expend the effort necessary to make cartridges compatible with Sega's console without copying. The underlying premise that led the Ninth Circuit to declare a “market” in Sega cartridges, and to encourage competition in it, was probably Sega's success in the home entertainment field.\(^{231}\)

While quite right from an antitrust perspective, this criticism does not advance the copyright analysis. From an antitrust perspective, a market defined to include only Sega-compatible games ignores

\(^{227}\) *Id.* at 1521-22 (quoting 17 U.S.C. § 107 (1994)).

\(^{228}\) Indeed, less than a year ago, news reports stated that Nintendo was “locked in a battle with Sega Enterprises, Ltd. and Sony Corp., both of which introduced players last September that use powerful computer chips to provide faster and more colorful games than Nintendo's older machines.” *Playing the Waiting Game, Nintendo Touts Its New Ultra 64 But Says It Won't Release It Until April, Dallas Morning News*, Nov. 27, 1995 at 2D; see also Laurence Zuckerman, *Sega's Chief Steps Down*, *Int'l Herald Tribune*, July 17, 1996 (Finance), available in Lexis, News Library (describing continued competition in the video-game player market).

\(^{229}\) *Sega*, 977 F.2d at 1523.

\(^{230}\) *Id.* at 1523-24.

\(^{231}\) Miller, *supra* note 71, at 1019-20 (citations omitted).
competition from Nintendo (and now Sony), and thus ignores the discipline such competition imposes on Sega's conduct with respect to its console. Among other things, such competition could provide Accolade with the opportunity to play the primary competitors off against one another, if Accolade's intellectual property was sufficiently desirable to prompt an auction. Although these arguments were rejected in *Kodak*, the discussion in the previous section demonstrates that they retain considerable force. Discipline in the primary market is relevant because it will keep Sega from doing the things antitrust seeks to prevent—reducing output to price above marginal cost or colluding with others to do so.

The antitrust analysis will not do for copyright purposes because market definition in the fair use context raises an important difference between antitrust and copyright. Copyright is not content with the conclusion that Sega may not price above marginal cost, and thus copyright analysis cannot end at the point this conclusion is reached. Copyright is concerned with promoting innovation, which increases the array of goods available to consumers. Were Sega a monopolist in a market with highly elastic supply, for example, antitrust would not be concerned, but copyright would. This difference is relevant to the market definition issue because innovation requires investment, which (and this is the entire premise of the Copyright Act) requires that copyright preserve enough of the market for the copyright holder to obtain the competitive return necessary to induce investment. It is thus not sufficient for copyright purposes to identify a market that constrains supracompetitive pricing; the relevant market must be related to the returns available to the copyright holder.

In cases involving aftermarkets, a proper approach to market definition will necessarily entail consideration of the effect of the defense on all markets from which the returns on copyrighted material are derived because rational firms will calculate returns from all such markets in deciding whether to invest in intellectual property. This is particularly important in the software context. From the consumer's point of view, there is no demand for Sega games apart from Sega consoles, nor for consoles without games. Consumers of Sega's products want to play video games. This requires a system entailing both hardware and software, which are therefore extremely strong complementary goods. Sega is in this respect similar to *Kodak*; copier purchasers did not want machines, parts, or service as such—they wanted copies. This does not mean that the aftermarkets may not constitute legitimate markets for antitrust purposes. On this point
**Kodak** is right, although as noted above this conclusion in and of itself cannot justify the invocation of antitrust in aftermarket cases. When combined with copyright's concern for returns to the copyright holder, however, it means that antitrust's narrow focus on supracompetitive pricing in these aftermarketss is inadequate for copyright purposes, and thus that these markets may not be viewed in isolation from the primary market, as the **Kodak** court did. The court in **Kodak** was in part concerned that Kodak would shift its source of revenues from the competitive primary markets into the uncompetitive aftermarketss, raising among other things the possibility of price discrimination in the aftermarketss. Because copyright is concerned with returns to the copyright holder, such practices are not necessarily evil. To the contrary, copyright fulfills its mission precisely by giving whatever degree of market "power" copyright holders may enjoy through a prohibition on copying.

In this respect, the opinion in **Sega** is analytically deficient. The fourth fair use factor asks what portion of the market will be put at risk by the defendant's use; the broader the market definition, the smaller the amount of that market that could be "usurped," and the more likely a fair use defense would be to succeed. The **Sega** court was concerned with justifying its fair use defense and, as we shall see in a moment, with establishing that little of the relevant market would be

---

232. There are differences: copier purchasers could perhaps reduce their consumption of service (though perhaps not parts) by stretching the period between service calls, while playing a video game always requires both a game and a console. Moreover, Accolade might well be able to build a hardware platform for its games; it would not have made sense from any perspective (even had it been possible, which it almost assuredly was not) for the ISOs in **Kodak** to go into the copier business.

233. As noted above, this analysis was improper in **Kodak** as an antitrust matter; it is improper in the copyright context both for those reasons and for the additional reasons unique to copyright's concern with economic returns to copyright holders discussed in the text.

234. **Kodak**, 504 U.S. at 480-86.

235. Considering all markets relevant to an integrated system such as computer games does not predictably favor either copyright holders or alleged infringers asserting a fair use defense. The broader the market definition the narrower the share any given firm is considered to have. Moreover, an increase in the amount of software that will run on a given item of hardware will confer some benefit on those who reap returns from the hardware (a phenomenon examined formally under the proper rubric of network effects below), as the well-known fight between Beta and VHS standards for video recorders makes clear. Entry of a new source of software (such as Accolade) for a given item of hardware (such as Sega's console) may benefit the hardware manufacturer overall; it may be that the availability of Accolade games would make Sega's consoles more popular, thus giving Sega a broader (though competitive) market for its games. Examination of both the primary market and aftermarket will capture any such benefits, which an examination of the aftermarket alone would miss.

236. **Sega**, 977 F.2d at 1523-24.
The court thus presumably would have been interested in a broad market definition; in this case one keyed to the primary market, which would have resulted in the conclusion that less of the market would be preempted (Accolade's share presumably being a constant). Instead, the court opted, without discussion, for the narrower aftermarket, which it characterized with its "monopoly" rhetoric. The court could only have meant that Sega was seeking a monopoly in games that played on its (non-monopolistic) consoles; the consoles themselves competed in the primary market against Nintendo. Unless Sega began manufacturing games that worked on Nintendo consoles, which for reasons discussed in the following section it would have every reason not to do, Sega's returns on its copyrights were keyed to this market as well: its games could be no more popular than its consoles. In short, both markets were relevant to the returns Sega would realize on its intellectual property. Therefore, the effect of Accolade's use on both markets should have been evaluated.

The Sega court's logical slip (not compelled by the Copyright Act itself) is revealing. The court was not thinking of markets in terms of returns to copyright holders; instead it sought to support the broader purpose of expanding the variety of available computer games. The rhetoric of "monopoly" was useful for this purpose: it implied a need for a ruling that would "bust up" the monopoly and allow new competitors into the "market." Still, it is hard to know how useful the loaded term "monopoly" can be in these circumstances. It did not mean that Sega could raise the prices of either its consoles or its games in a monopolistic manner—Nintendo, later joined by Sony, prevented that from happening, and in any event copyright would not necessarily be offended if Sega earned supracompetitive returns. Nor could it meaningfully refer to Sega's dominance of the market for games that operate on its console. Similar to Kodak's position in parts and service for its copiers, Sega is likely to have a large, if not dominant, position in the "market" for games that run on its consoles because it created the market. If nothing else, Sega's costs will be lower because it does not have to reverse engineer its own games, meaning that it can alter the interfaces of its consoles more cheaply than its competitors can reverse engineer them.

237. Id. at 1523.
238. Id. at 1526.
Particularly if consumers are able to buy more than one hardware system (an option evidently unnecessary and too expensive in Kodak but perhaps not uncommon with respect to computer games), a firm's dominance in the aftermarket for games that play on its consoles does not imply a paucity of computer game options for consumers, any more than it implied in Kodak a paucity of options for those who wanted to purchase copiers. There might well be more programs available if Accolade could copy Sega's programs in some circumstances, but that is a truism that does not help draw the line between permissible and impermissible copying and has nothing to do with the accepted meaning of "monopoly." There would be more games if all console manufacturers were taxed and the money sent to subsidize third party game developers, but that does not mean that it would be a good idea. Variety in intellectual property may be the primary goal of the copyright laws, but it addresses only one side of the tension between promoting creativity and preserving competitive returns; that tension simply cannot be resolved by ignoring the other half of the equation. The Sega court's failure to confront the market definition issue, and its reliance on Sega's share of the aftermarket it defined, truncated the analysis at precisely the point it began to get interesting and weakened the court's hasty conclusion that Sega's returns would be adequately preserved.

However the market was defined, the Sega court's dichotomy between usurpation and competition within a market is neither a tenable interpretation of Harper & Row nor a workable approach to evaluating market preemption in the fair use context. Harper &

240. This analysis, of course, threatens to become circular at one level: Sega has this cost advantage only because of the copyright laws, which could be changed to eliminate that advantage, and thus perhaps promote competition, if this were deemed desirable. The proper analysis for evaluating such a proposal is discussed below.


242. The usurpation standard emerges from Harper & Row, which involved the unauthorized printing by The Nation of verbatim excerpts of President Ford's memoirs in an effort to "scoop" a Time magazine review of the book. Time, which had contracted for the exclusive right to publish and review excerpts of the book first, cancelled its article and refused to pay the last installment on its contract. On these facts, the Court concluded that the effective usurpation of a "market" in and of itself negated a fair use defense. Harper & Row, 471 U.S. at 567-69. Contrary to Professor Cohen's construction, no markets were "usurped" in anything approaching a literal sense. Cohen, supra note 150, at 1127. The book still sold some copies, and presumably was reviewed in other places. The publisher, however, lost $12,500 on the Time contract, and whatever amount of sales that would have been generated by a cover article in Time but were not generated by the article in the Nation. Harper & Row, 471 U.S. at 567. The
Row held that "a use that supplants any part of the normal market for a copyrighted work would ordinarily be considered an infringement" and that "isolated instances of minor infringements, when multiplied many times, become in the aggregate a major inroad on copyright that must be prevented." The Sega court thus appears to have broadened the scope of fair use beyond the point approved by Harper & Row by allowing a copier to seize a greater share of relevant sales than previously was found to defeat a fair use defense. Moreover, the question whether investment in intellectual property will be undertaken depends on the rate of return to be earned by the investment. The calculation of such returns ex ante does not ask simply whether a market is likely to exist or be completely usurped, because markets themselves are not all-or-nothing creatures. Even a "monopolistic" market in Sega's terms may not induce investment if competition from substitute works is severe, and returns on software might be sufficient to induce investment even if the Accolades of the world were allowed to seize large portions of the relevant market.

Perhaps there is another way to look at the situation; one that does not exceed Harper & Row's boundaries. The Ninth Circuit's analysis of the first factor concluded that Accolade's computer games were comprised of Accolade's own intellectual property. Accolade was not copying Sega's property to help Accolade develop a competitive game, it was copying Sega's copyrightable code only to study the functional, noncopyrightable code, which it then adapted to allow the game it already had developed to compete. As relevant to the fourth fair use factor, this conclusion implies that Accolade's "use" of Sega's copyrightable material had no market effect at all—Accolade "used" that material only to the extent that it identified it as

---

245. See BREALEY & MYERS, supra note 37, at 13-14; see infra text accompanying note 267.
246. Sega, 977 F.2d at 1522.
247. Id.
expressive code, rather than the functional interface code Accolade needed, and then moved on. Because Accolade only incorporated the adaptation of the nonprotected functional elements of Sega's interface code into its program, any adverse effect on Sega's sales would come from a consumer preference for Accolade's own intellectual property over Sega's. As the court said, "it is the characteristics of the game program as experienced by the user that determine the program's commercial success . . . . [T]here is nothing in the record that suggests that Accolade copied any of those elements." This is exceedingly powerful reasoning: carried to its logical conclusion it holds that Accolade could have usurped Sega's market completely without vitiating its fair use defense because the diminution in Sega's market share would not stem from Accolade's infringing "use" of copyrightable material, which is the inquiry dictated by the fourth factor.

Against this Sega has a simple response, which is not without some resonance. But for its infringement, Accolade would not be competing in Sega's aftermarket at all. And if we conclude that copyright would permit Accolade to drive Sega from the aftermarket for Sega's consoles, Sega will effectively be reduced to the status of a manufacturer of video recorders. There may be nothing wrong with this. Many firms make video recorders and are perfectly happy doing it, and from an antitrust perspective we should say that if consumers prefer Accolade's product to Sega's, then Sega deserves to be driven from the market. But can we afford to be so sanguine? The market evidence is that Sega does not want to be just a console manufacturer, it wants to provide consumers with a video game entertainment system, entailing both the console and the software. Indeed, given what it likely thought were its rights under the copyright laws, Sega may have sold consoles at or near cost, planning on earning most of its return through the sale of compatible software. From an ex ante perspective, Sega might well not have invested in either hardware or

248. *Id.* at 1523.

249. The strength of the argument is analogous to the strength of the related argument with respect to MAI and Data General: because Accolade presumably would compete on the merits of its own intellectual property rather than through the use of Sega's (copyrightable) intellectual property, Sega's infringement claim may appear to be a tactical use of copyright to obtain a competitive advantage beyond that which the copyright laws are intended to confer.

software if it feared it could be driven from the software aftermarket through reverse engineering such as Accolade's. If that response held for firms generally, society would suffer a net loss of intellectual property: without Sega's consoles, there is no aftermarket for Accolade to enter.\footnote{This introduces yet another complex issue. If Congress or the Supreme Court were simply to announce that henceforth reverse engineering would be allowed and firms such as Sega would have to compete in their aftermarkets as well as with Nintendo and Sony, then firms such as Sega would take this into account in deciding whether to invest in computer game systems. If the returns were deemed inadequate, they would not invest. There is nothing unfair about that. Firms make investment decisions based on the legal regime all the time, as the frequent use of the Internal Revenue Code to induce or penalize investment demonstrates. Sega might well protest that its investment is already made, and the legal tables were turned on it \textit{ex post}, thus essentially casting itself in the role of purchasers of Kodak copiers prior to the implementation of Kodak's restrictive parts policies. This point is at best partially true; not only do firms make investment decisions based on the legal regime, it is also no secret that the regime may change in midstream, and smart firms should factor such uncertainty into their decisionmaking as well. Still, the court in \textit{Sega} was quite candid that its result might seem surprising at first blush. \cite{Sega_1994} at 1527. The case would not have occasioned the amount of commentary it has received had it been easily anticipated from previous cases. \textit{See}, e.g., Victor de Gyarfas, Note and Comment, \textit{Sega v. Accolade: A Step Forward for Reverse Engineering}? \textit{23 Sw. U. L. Rev.} \texttt{571} (1994); William S. Coats and Heather D. Rafter, \textit{The Games People Play: Sega v. Accolade and the Right to Reverse Engineer Software}, \textit{15 Hastings Comm/Ent L.J.} \texttt{557} (Spring 1993); Christopher W. Hager, \textit{Apples and Oranges: Reverse Engineering as a Fair Use After Atari v. Nintendo and Sega v. Accolade}, \textit{20 Rutgers Computer \\& Tech. L.J.} \texttt{259} (1994).}

There is logical force to both arguments, but precious little empirical basis for untangling the competing logic. In such circumstances, the best courts can do is to identify all the relevant factors and possible consequences of ruling one way or the other, examine such evidence as the market has to offer with respect to each factor, and proceed cautiously as far as the evidence warrants, but not one step farther. The basic premise of copyright rightly implies that the scope of protection a copyright affords correlates with the returns flowing to copyright holders. One cannot adjust the scope of protection without a corresponding effect on the returns. It follows that allowing reverse engineering will likely result in the loss of some software that would have been produced under a more protective regime. This does \textit{not} mean that the \textit{Sega} court was wrong, however. It may well be that the marginal loss is insignificant because if Sega faced no competition in its aftermarket, the additional software it produced might be complacent and uninteresting. That is highly unlikely of course, because Sega has to compete with Nintendo and Sony, but in other circumstances it might be a valid consideration. Alternatively, the increase in software induced by a regime that allowed reverse
engineering might more than make up for the marginal loss. This reasoning seems to underlie the Sega court’s assertion that allowing Accolade to compete would not materially reduce the market for Sega’s software. Either way, however, we cannot accept Sega’s implicit conclusion, which Professor Cohen aptly summarizes as being that “the purposes of copyright—to encourage the production and distribution of creative works—is best served by allowing such competition, not by blocking new market entrants,” without a more detailed analysis of the actual processes by which firms within the software industry make investment decisions.

Unfortunately, the Ninth Circuit did not construct an appropriate analytical framework to deal with these complex issues. The fourth fair use factor cannot be sensibly applied without a standard for distinguishing cases in which an infringer seizes too much of the market to assert the defense and cases in which the defense could succeed. In other words, the real question in these cases is how much of the market can an infringer take before its use is unfair? Contrary to Professor Cohen’s reading, the Ninth Circuit never confronted this issue squarely. It merely argued that Harper & Row was distinguishable because, as Professor Cohen puts it, that case “involved a ‘scoop’ of the heart of a copyrighted work that threatened to supplant the market for the work entirely” and because consumers would not buy more than one copy of President Ford’s memoirs while they do buy multiple copies of video games.

It is hard to see how Harper & Row supports such a reading. The Court there rejected a fair use defense based upon articles written about President Ford’s book, rather than sales of the book (the copyrighted work) itself. There was no question whether the “good parts” of the book would be serialized in a magazine, the only question was whether they would be serialized by Time for a fee or by the Nation for free. Serialization of that particular book may have diminished rather than spurred sales; either way, however, that factor was a constant feature and thus could not sustain the rejection of the fair use defense.

252. Cohen, supra note 150, at 1128.
253. Professor Cohen agrees: “The question remains, as always, at what point—short of every use, or every ‘commercial’ use, of copyrighted material, to draw the line.” Cohen, supra note 150, at 1126.
254. Id. at 1127 (arguing that Sega directly addressed the materiality question).
255. Sega, 977 F.2d at 1523.
257. Id. at 543.
The court's assertion in *Sega* that consumers will buy Accolade's games as well as Sega's, and thus little if any preemption will occur even in the narrowly defined "Sega-games market," is similarly problematic. According to the court, it does not "seem unlikely that a consumer particularly interested in sports might purchase both Accolade's 'Mike Ditka Power Football' and Sega's 'Joe Montana Football,' particularly if the games are, as Accolade contends, not substantially similar." But if the games are not substantially similar in any meaningful sense, they presumably are not substitutes, not competitors, and thus irrelevant to the market displacement issue, though they would be relevant to Sega's potential market. This is perhaps a minor detail, because the substantial similarity argument is untenable either way. If the games are identical they will compete directly for sales to consumers with a taste for Sega-compatible football, baseball, or whatever type of games. Even if they are not, so long as consumers are unwilling to spend infinite sums on Sega-compatible games, Accolade's entry with any type of game, identical or not, will present consumers with alternatives to Sega and Sega-licensed games for their Sega-compatible dollars, thus diminishing Sega's potential and actual sales.

As noted, the court attempted to avoid the problem by positing multiple purchases of every game. But the proposition that consumers will purchase both Mike Ditka's football game and Joe Montana's football game is hardly so obvious that it is properly subject to judicial notice on appeal. One would at least expect an analysis of the incomes of game purchasers to determine whether the average purchaser could afford to add two of every game to their electronic libraries. No market is infinite. At some point the game would turn zero sum as it likely would be from the start. In any event *Harper & Row* cautions that the "potential market for the copyrighted work" is entitled to some protection, and that market includes sales that could go to Accolade if Accolade were to market Sega-compatible games.

258. *Sega*, 977 F.2d at 1523.
259. *Id.*
260. It is similarly not apparent why the court felt it proper to take Accolade's contentions at face value on an appeal from entry of a preliminary injunction after evidentiary hearings.
261. This is not necessarily to say that such relatively indirect market preemption would necessarily defeat a fair use defense; the point simply establishes that the Ninth Circuit's analysis was materially incomplete.
262. *Sega*, 977 F.2d at 1523.
Sega's analysis thus simply does not address the issue at the heart of the fourth fair use element.

The doctrinal label we are given to deal with the problem is "materiality." Harper & Row endorses Professor Nimmer's view that the fair use defense should be "limited to copying by others which does not materially impair the marketability of the work which is copied." While the Sega court perhaps did not apply the usurpation standard as stringently as a literal reading of Harper & Row might dictate, its approach may have satisfied the spirit of that opinion and of the fair use doctrine when the unique features of software are taken into account. Sega rightly based its analysis on copyright first principles, which also may be used to derive a rule of materiality for the fourth fair use factor, taking into account the economic characteristics of software markets. Applying the basic copyright assumption that innovation is a function of willingness to invest, probably the best feasible approach would be to draw a rough analogy to the materiality standards developed under the federal securities laws and ask, on a project-by-project basis within a given firm, whether the amount of the market foreclosed by reverse engineering would reduce the returns to a copyright holder by an amount that, ex ante, would deter an economically rational firm from undertaking a given investment in a particular software project.

This analysis, which must be undertaken with respect to the expected returns of each project (rather than with respect to the wealth of the firm as a whole), entails the familiar calculation of net present value: discounting the cash flows to be realized from the investment through an appropriate terminal point, using a discount rate accurately reflecting the firm's cost of capital (the amount the firm could earn by investing the funds in a security of equivalent risk to the project), less the cost of the investment. This rule places the fair use focus on the question whether a change in facts—here the legal protection afforded copyrights—would affect an investment decision, which is precisely the analysis needed for this factor. Choosing to permit reverse engineering, and thus to permit competition in the Sega aftermarket, will likely diminish the expected

264. Id. at 566-67 (quoting 1 Nimmer on Copyright § 1.10[D] at p. 1-87) (emphasis added).
266. For a more detailed discussion of the process by which project net present values are calculated, see BREALEY & MYERS, supra note 37, at 12-15.
return on Sega’s investment relative to the returns that could be had in a regime in which reverse engineering was not fair use. By increasing the risk, Sega would not realize the number of sales it would under the more protective regime. This in turn reduces the discounted cash flow figures (either through a reduction in the number of sales projected or, what is the same thing, an increase in the estimated risk that the original sales figures would not be reached). The willingness to invest in intellectual property, relative to a more protective regime, likewise would be reduced.

The conceptual rule derived from this standard is easy to state: If Accolade’s copying facilitates the foreclosure of enough of Sega’s sales and potential sales that Sega would not have invested in either its game console system or any given video game *ex ante*, then Accolade’s fair use defense should fail as a matter of law, regardless of the outcome of the other three fair use factors. If, as may well be the case, Accolade’s use had a positive effect on demand for Sega’s consoles, this effect would go into the mix in Accolade’s favor; depending on the facts of any given case, such positive feedback effects might well be conclusive. Because Sega’s revenues from the system are constrained by the popularity of the consoles, an increase in the popularity of the consoles would at a minimum increase Sega’s potential revenues. If a *de minimis* portion of Sega’s expected returns would be lost, this factor would weigh in Accolade’s favor. At the point the copyright holder has been assured sufficient returns to induce investment in intellectual property, copyright’s utilitarian focus is satisfied and its principal concern for enhancing the variety of intellectual property available in the market should take precedence. If, as might often likely be the case, no such clear determination could be made, the factor should tip in favor of the copyright holder on the ground that the costs of erroneous decisions are asymmetric. It is more important to secure returns for those who create technology in the first instance, and thus to secure the creation of the technology from which further innovation will spring, than it is to secure the right to copy that technology once it is made. If mistakes create disincentives to invest in intellectual property, issues regarding the fair use of such will be moot.

One of course cannot pretend that this rule can be precisely applied with ease in actual cases. There are numerous practical difficulties with such a rule that will vary according to firms and industries. Firms may have different “hurdle rates” for investment; of two rational firms one might automatically accept a positive net present value investment (assuming the availability of capital through
financing or retained earnings), while another might reject the same investment in the hope that an investment with a higher net present value could be found. Even without such relatively idiosyncratic factors, arriving at a good estimate of the relevant variables would be extremely difficult. And as objective truth about such matters simply does not exist outside the models of financial economics, calculations of one side’s experts could readily be challenged by an opponent’s experts, thus increasing the costs of litigation while arguably doing little to enhance the accuracy of the ultimate determination.

Nevertheless, those shortcomings are endemic to any serious analysis of the question; they are a product of the decision to resolve such issues through a judicial process rather than some other. Any serious materiality analysis, implicit or explicit, makes assumptions about firms’ investment behavior, whether those assumptions are acknowledged in the analysis or not. As discussed above, Sega too hastily assumed that Accolade would not preempt much of Sega’s aftermarket, but it also assumed, without any discussion, that a small reduction in Sega’s share of its aftermarket would not affect its willingness to invest in intellectual property. From an ex ante perspective, that may or may not be true, though with the high fixed and low variable costs characteristic of software it is perhaps a safe assumption. But we should not be satisfied with guesses and assumptions when analytical tools are available to help. Just as it will

267. At a minimum, one would need to know the amount investors expected the firm to make, discounted by the probability of failure, and the return available on investments of similar risk. None of these figures is easily derived, however, and seemingly small alterations in assumptions can produce large differences in value. For example, assume a firm undertakes to develop a program that will require an initial investment of $1 million, a $300,000 investment at the end of the first year of development, and a $300,000 investment at the end of the second year; assuming a discount rate (cost of capital) of 15%, the present value of the cost of the investment is $1,487,800: the $1 million immediate cost plus ($300,000 x .870) + ($300,000 x .756) (The multipliers in this note are taken from standard present value and annuity tables.). Assume further that the software will generate $650,000 in revenue for four years beginning at the end of year 2; the revenue over the expected life of the project is thus $1,859,000 ($650,000 x 2.86). As this is the value after two years, it must be discounted ($1,859,000 x .756), yielding a present value of revenues of $1,405,404 and a net present value of ($82,396). This investment is a money-loser, and will not be undertaken by a rational firm. If we alter the risk associated with the project, however, such as by increasing the scope of the copyright, a different result may be obtained. Assume that, due to an increase in protection from copyright, the proper discount rate for the investment is 12%. Under that assumption the cost is $1,507,000 and the revenues are $1,976,000, discounted to $1,574,872. The present value of the investment under these assumptions is $67,872; this investment will be undertaken and copyright’s purpose of promoting innovation will be achieved.

268. Sega, 977 F.2d at 1523.
not do simply to assume that only a small portion of the market will be foreclosed, so it will not do simply to assume that whatever portion of the market is foreclosed will or will not be significant in terms of the copyright laws. Firms are not mysterious black boxes whose decisions may simply be assumed to conform to the dictates of copyright or any other area of substantive law. Those decisions must be analyzed in light of the internal economic processes by which the decisions are made.

2. Reverse Engineering, Replication, And Essential Facilities

What, the patient reader may justifiably ask, has all this to do with the essential facilities doctrine? The answer is that, particularly in Kodak-style aftermarkets, the scope of the fair use defense in the reverse engineering context is a key to making sense of the question whether the "facility" (in these cases the software interface) is capable of being replicated. This in turn is a key to whether the natural monopoly conditions necessary for sensible application of the essential facilities doctrine are present in a given case. This question was not presented as such in Sega, but the analytical approach taken in that case is necessary to resolve it. To see why, we may simply add an antitrust counterclaim to Accolade's defense of Sega's infringement action: Assume Accolade alleges that Sega is attempting to exclude it from the aftermarket for Sega-compatible software, and thus raises an essential facilities claim in an effort to gain compulsory access to Sega's software.

The addition of a counterclaim presents an important definitional issue, similar to the question whether MAI's diagnostic software could be considered a facility apart from MAI's computers: what, exactly, is the "facility" to which Accolade would demand access? Sega's console is not an input to Accolade, any more than a car is an input to a tire manufacturer; the console is a complementary good necessary for the software to satisfy consumer demand for a video game system, and vice versa. Moreover, the market evidence showing three firms actively competing with different consoles demonstrates conclusively that consoles themselves do not possess natural monopoly characteristics. Accolade's claim thus would necessarily rest on the narrower theory that the ability to interface with Sega's console was a facility essential for Accolade to compete in the market for Sega-compatible computer games. Absent the ability to reverse engineer software, Accolade would say, it could satisfy one element of a valid essential facilities claim—the interface would not be replicable at any
cost. Though this would be a consequence of copyright’s grant of protection to Sega, rather than natural monopoly, the competitive effect would be the same.

This modification to Accolade’s defenses cleanly presents the question at the heart of the antitrust/copyright debate: may compatibility—the ability for two systems to interface—count as a “facility” for antitrust purposes? Unlike most of the traditional antitrust essential facilities cases, compatibility is of course not a separate input; it is a feature of two systems rather than a product or service itself. The distinction between inputs and compatibility is unstable. Some programs are sold whose sole purpose is to translate data from one software format to another—allowing a file written on WordPerfect, for example, to be stored and edited in a format that Microsoft Word could use. Under the Jefferson Parish approach\(^{269}\) it would follow that a “market” for interface code exists, and this would tend to drive the analysis back toward more familiar ground. But we should not allow the definitional question to distract us from the larger point. Consumers have no desire for video game software as such, they want to play video games, which is to say they desire the product of the complementary relationship between the video game and the console. A software producer such as Accolade thus always has the option of building its own consoles for its own games. While this will not get Accolade into the Sega aftermarket, it will allow Accolade to sell a complete video game entertainment system, which is what consumers are buying. Assuming (as Accolade assured the Ninth Circuit was the case) that Accolade’s games are distinctive, Accolade’s software will likely stand on its own regardless whether it is on a Nintendo, Sega, or Accolade system. Copyright’s interest in the production of software would thus be satisfied, though at some cost. On the other hand, if the cost of switching from a Sega console to an Accolade console, or of buying one of each type of console, is prohibitive, and if reverse engineering were permitted, one would expect separate interface software to be designed and sold.

No matter how technology advances, the antitrust analysis remains the same. A facility is replicable, and thus not essential insofar as the antitrust laws are concerned, so long as reverse engineering is permitted as fair use.\(^{270}\) If a firm can reverse engineer software at a


\(^{270}\) This of course includes the assumptions that no copyrightable code is incorporated into the work of the firm engaging in the reverse engineering and that, as was the case in Sega, only
cost that allows it to make a profit from an aftermarket, it will do so, as Accolade did. If a firm cannot do so, it will not. Whether reverse engineering is profitable will be a function of the complexity of the reverse engineering task and the potential size of the aftermarket (along with the constant concerns of the risk associated with projected returns and the existence of possibly more remunerative projects in which to invest). Some software aftermarket markets may be so small that the returns they offer will not justify investment in reverse engineering. It does not follow, however, that with respect to such markets compatibility becomes an essential facility. By hypothesis the market in such cases offered an adequate return for the firm that created the market; that the market will not support both that firm and an aftermarket competitor does not make it a natural monopoly market. In most conceivable cases, an aftermarket software producer will have the option of building its own platform for its software and competing in the primary market. As noted above, the costs incurred to enter primary markets are not a warrant for invocation of antitrust; they are simply a fact of business life that cannot be changed without serious anticompetitive distributional consequences.

It is important to stress here that the foregoing analysis does not imply that without the ability to reverse engineer software code, antitrust would have an appropriate role to play in software markets. If firms could not copy code for the purpose of achieving interoperability, the "facility" of interoperability would not be replicable. But this would be a consequence of the strength of the copyright, and thus would not properly be redressed through the invocation of antitrust. If Congress wants to allow copyrightholders to preclude any copying at all, in order to extract revenues from both the expressive elements of a copyrighted work and through the ability to interface with that work, that is Congress’ business. As noted at length above, antitrust has no ability to determine whether a copyright is too strong, and thus cannot properly be invoked to tinker with the parameters of the copyright.

---

the noncopyrightable interface code would be adapted into the firm’s own software. Any use of copyrightable code would fall outside this analysis; because use of copyrightable code would be unnecessary to achieve compatibility, which by hypothesis is the “facility” in question, the reverse engineer would have no recourse under antitrust principles to justify its conduct, which would be judged solely under copyright standards.
C. Strategic Contractual Behavior And The Essential Facilities Doctrine

Should that be the end of the analysis? Could a firm ever invoke the antitrust laws to argue that it should not be required to engage in reverse engineering, assuming the law permitted it to do so and that it would do so in compliance with the law? Or should a defendant that engages in anticompetitive strategic behavior be penalized under antitrust generally, or under the essential facilities doctrine in particular, or both? Take the claim, asserted in Data General\textsuperscript{271} (and Olympia Leasing),\textsuperscript{272} that the creator of a system that results in an aftermarket (such as a minicomputer or, less plausibly, telephone systems sales), carries an obligation to maintain the aftermarket for firms that choose to enter it. Data General’s experience with ISOs prior to the Grumman suit had been varied.\textsuperscript{273} When it began manufacturing computers in the 1970s, Data General had been wary of ISOs, which were often comprised of former Data General employees.\textsuperscript{274} Data General feared ISO firms would violate Data General’s intellectual property rights or misappropriate its trade secrets.\textsuperscript{275} At one point Data General filed suit against an ISO, which resulted in a settlement in which the ISO agreed to return to Data General any proprietary information certain former Data General employees had taken, and Data General agreed to authorize the ISO to use Data General’s proprietary information in the future.\textsuperscript{276} For the next several years, Data General followed a relatively liberal policy: selling or licensing its diagnostics directly to ISOs; allowing ISOs to use diagnostics belonging to the owners of its computers when the ISOs worked on them; allowing ISO technicians to attend Data General training; and even repairing some circuit boards at the request of ISOs.\textsuperscript{277} This more liberal policy lasted until the mid-1980s, when Data General began refusing to allow ISOs to purchase its diagnostic devices, refusing to repair parts for ISO customers, and refusing to allow ISO technicians to attend Data General training.\textsuperscript{278}

\textsuperscript{271} Data General, 36 F.3d at 1188.
\textsuperscript{272} Olympia Equip. Leasing Co. v. Western Union Tel. Co., 797 F.2d 370, 379 (7th Cir. 1986).
\textsuperscript{273} Data General, 36 F.3d at 1152-54.
\textsuperscript{274} Id. at 1155-54.
\textsuperscript{275} Id. at 1153.
\textsuperscript{276} Id. at 1155.
\textsuperscript{277} Id. at 1153-54.
\textsuperscript{278} Id. at 1154. One could perhaps correlate the timing of Data General's change of heart on the ISO question with the beginning of the end of the minicomputer market and the
As noted briefly above, Grumman contended that *Aspen Skiing* established that a "monopolist that has helped a market to develop may not withdraw its support without legitimate business justifications." The First Circuit viewed it somewhat differently, stating the problem presented by Grumman’s defense as "whether (and to what extent) the antitrust laws, in the absence of any statutory exemption, must tolerate short-term harm to the competitive process when such harm is caused by the otherwise lawful exercise of an economically potent ‘monopoly’ in a copyrighted work.” The two questions are related, however, if one recasts Grumman’s argument slightly to say that Data General induced Grumman’s investment in the service market in order to strengthen its own position in the primary market. This case would then present issues of opportunism not presented by MAI and, contrary to Professor Fox’s contention, not presented by *Kodak* either. Thus recast, the argument presents the possibility of suboptimal investment, and harm to consumer welfare, based upon conduct by Data General and reciprocal conduct by Grumman amounting to a tacit agreement, in the economic if not legal ascendance of powerful personal computers. The ISOs have plausibly claimed that, faced with a doomed market, Data General and other firms in its position cast about for ways to wrest the last drop of revenue from the systems they had in place. One way to do this was through supracompetitive pricing in the service aftermarket, in which Data General could profitably charge just under its customers’ switching costs. This is of course the end-period problem discussed above in connection with *Kodak*. Though real, the problem presents mostly distributional issues between Data General and its customers, and Data General still faces certain reputational constraints to the extent it hopes to be a repeat player in different market segments, though the constraints are admittedly weakened by the demise of the minicomputer market. The previous discussion regarding the advisability of relying on such factors as a warrant for invoking antitrust applies here as well.

279. *Id.* at 1188. Though the First Circuit rejected this interpretation, its reasoning elided more of *Aspen Skiing* than it explained. We thus cannot be sure that *Aspen Skiing* will not lead to more mischief in the First Circuit, to say nothing of the rest of the country.

280. *Data General*, 36 F.3d at 1152.

281. Professor Fox at one point casts the issue decided in *Kodak* as concerning “a cut-off of well-performing independents by the firm that had induced their investment in the market.” Fox, *supra* note 105, at 761 (emphasis added). But there is no evidence of inducement in the ordinary sense in the *Kodak* opinion itself. At most, the ISOs could argue that Kodak induced their investment by developing its copiers and thus their aftermarket. But this would seem to better support Kodak’s free-rider defense to its policy than any allegations of wrongful behavior. See generally *Kodak*, 504 U.S. at 451. It would be uniquely perverse, even for antitrust, to hold that the development of a new product constituted some sort of attractive nuisance to firms wishing to engage in derivative activities. The situation might be different if the parties had some sort of agreement, such as in the cash register case to which Professor Fox analogized *Kodak*, Fox, *supra* note 105, at 761-67, but likely not. In that case the parties’ rights should be defined by their contract; a party’s failure to ensure a continued supply of parts in the contract would not suffice to create an antitrust violation, particularly where other parties stepped in to fill the void.
sense, subsequently breached by Data General without payment of damages to Grumman.282

Let us examine the components of this claim, and in the process move it from the mixed hardware/software context of ISO claims (which in any event will fade in time as the installed base expires) into a wholly software environment. For Data General, substitute the firm in our second hypothetical, which makes an application program for home banking ("A"). For Grumman, substitute a software firm that does not have a home banking program, but does have a program that will calculate foreign currency hedges and do other similarly impressive things ("C"). When A enters the market it will be concerned, like all firms, with expanding its share as rapidly as possible. It will also have an eye towards its competitors' products, and will seek ways to make its software more desirable. Assuming there is a market for home currency hedges, the possibility of having C's program work with A's program will be enticing to A; and as C has no program of its own, the possibility for mutual benefit is obvious. Suppose A and C agree that A will provide C with the code necessary to interface with A's program for five years, and C agrees that it will write the interface into its program and aggressively market the program for the length of the contract.

Suppose that after 18 months A is established as the standard home banking program (with, say, 75% of the market), in part because consumers desired to hedge currency at home and thus valued C's compatibility with A. Suppose further that A could maintain 60% of the market even without C, but that C has no immediately viable alternative to compatibility with A. Lastly, suppose that A informs C that the original agreement covered only version A.1 of the home banking program, that A will be coming out with a new version A.1.1, which adds cosmetic bells and whistles to the program, and that A offers to enter into an agreement to provide C with the necessary interface code if C will sign an agreement including the same terms as the original agreement plus a license fee of 30% of C's profits on each program sold. Assume for the moment that A's interpretation of the original agreement is both wrong and asserted in bad faith. Does antitrust have a role to play here?

282. Payment of damages is important in this context not because breach of the agreement would entail the transfer of wealth between firms; as we saw in discussing Kodak such a transfer is unobjectionable in and of itself, at least on economic grounds. But failure to enforce agreements would increase uncertainty and thus risk and thus deter parties from embarking on ventures that would otherwise be profitable under a regime of contractual enforcement.
The argument in favor of some form of relief for C is fairly straightforward. A induced C to commit resources over an identified period of time to the creation of a product the existence of which conferred a material benefit on A. Having obtained the benefit of increased popularity (market share) of its program, A then switched courses before the period was even halfway over and, using the power it had obtained in part through C’s performance of the agreement, held up C in an effort to extort more favorable terms. The extortion arises because C must have compatibility with A to sell its program, and to sell updates of its program to its installed-base owners. Either reverse engineering A’s code in order to achieve compatibility or writing its own home-banking software (including such code as was necessary to allow owners of A’s software to transition smoothly to C’s new program), would take time and money, which C might have a hard time getting if it could not sell programs compatible with the newest version of A. While such options might be viable in the long run, as C could presumably obtain financing if it could make money through either option, C could have a serious and perhaps fatal short-run problem.

Under such circumstances, could compatibility with A’s software be considered an essential facility for C, such that C could bring suit under the antitrust laws to compel A to provide C with the necessary code? The preceding analysis compels the conclusion that C could not. First, in such a situation A’s opportunistic behavior would be the only element of legitimate legal concern. Opportunism here has two components: inducement and alteration. Inducement is necessary to distinguish this hypothetical from cases in which a firm such as C independently developed software compatible with A’s software without dealing with A. A could bear no responsibility for such a decision, and thus could not bear responsibility for any incompatibility that might be caused by upgrades to A’s original program. In such a situation C will have assumed the risk that A might change its program. C’s decision to commit its resources to A-compatible software will reflect that risk, and thus presumably would not reflect a suboptimal investment of resources. Where A induced the investment with certain promises, however, the possibility of

283. C’s position in such a case would be no different from that of a lunch counter that opened near a factory to provide food for the workers. Absent some agreement with the factory, the owners of the lunch counter will have assumed the risk that the factory will close, and their business will go with it.
suboptimal investment exists. A’s opportunistic alteration of the deal, if unpunished, would increase the risk of such agreements and thus decrease the likelihood that they would be formed, resulting in harm to social welfare to the extent such agreements are desirable.

While relief for C is plainly in order, it does not follow that the relief should come in an antitrust package. Following Kodak we might say that a cognizable antitrust “market” for A-compatible software exists; that A has power in this market; and that compatibility with A is necessary for C to sell its software. We might further say that, at least in the short run, it is not economically feasible for C to replicate the interface with A by, for example, reverse engineering, or to write its own platform program and thus move into the applications market to compete against A. Such a sequence of reasoning, which could imply that recourse to the essential facilities doctrine would be proper in such cases, does more to show the perils of the Court’s analysis in Kodak than anything else.

If we conclude that C has no legitimate legal claim to A’s code beyond the period of the agreement, it follows that C is only harmed by the premature termination of the agreement. And should A deprive C of the benefits of the contract, the proper remedy would be for breach of the agreement, not for violation of the antitrust laws. Enforcement of the agreement according to its terms would give C everything to which it was entitled; antitrust would give it more, and indeed, would give C too much. The imposition of treble damages and attorney’s fees would interfere with the efficient operation of the contract by penalizing what might otherwise be an efficient breach by A. The risk that merely signing a contract could lead to trouble with the antitrust laws could reduce the willingness of A firms to sign such agreements, posing a risk of suboptimal investment that is the rough reciprocal of the risk posed by not enforcing the agreement at all. As noted above, antitrust has no role to play in refereeing contractual disputes. Allowing antitrust claims to proceed in such cases will merely distort the surrounding legal landscape and impede the operation of doctrines better suited to resolving the cases.

Approximately the same analysis compels us to reject the invocation of the antitrust laws in the intermediate case presented by Data General—in which the two firms have no formal contract, but in which one firm tolerates and perhaps encourages another to enter a market because such entry confers a benefit upon both the entrant and

---

284. This would be particularly true if A surprised C with its plans, without advance warning.
the existing firm. The contractual relationship of the firms in such a case may be ambiguous, ranging from no agreement at all to an implied undertaking by a Data General not to prejudice a Grumman once it had entered the market. The outcome of such cases, however, must be determined by where they fall on the contractual continuum, not by antitrust. If there is no agreement, the risks of suboptimal investment are low (a firm will presumably account for the risks it assumes), and there is no warrant for an antitrust claim. If there is an implied agreement, the parties should seek enforcement of the agreement, not file a monopolization claim. The claim that Aspen Skiing allows antitrust to be brought to bear in such cases is simply wrong.

D. Network Effects and the “Tipping” Argument for the Essential Facilities Doctrine

Computer software possesses certain economic characteristics that present difficult issues at the intersection of law and economics. These include the tendency of software to improve rapidly, thus mootng older versions and the statutory period of copyright, and the high fixed and low variable (and thus declining average) costs of producing software. Software also is characterized by “network effects,” which means that the value of software to a given purchaser increases as additional consumers buy the software as well.285 The more popular an application program is, for example, the more likely it is that others with whom a consumer deals will have or be able to use the program, making it easier to share data. Similarly, firms that produce complementary software will be more willing to develop such programs for popular application programs than for unpopular programs for the simple reason that popular programs will support more sales. Professors Katz and Shapiro usefully distinguish between direct and indirect network effects.286 The former term denotes networks that are physically linked or direct interactions using the same system; examples include telephones, fax machines, and language.287 The latter term denotes complementary relationships between products or simple demand-side economies of scale; the more people who own a particular brand of car or operating system, for

285. For excellent and only slightly technical discussions of this concept, see Katz and Shapiro, supra note 18; Stanley M. Besen & Joseph Farrell, Choosing How To Compete: Strategies and Tactics in Standardization, 8 J. Econ. PERSPECTIVES 117 (1994).
286. Katz and Shapiro, Systems Competition, supra note 250.
287. Id.
example, the easier it will be for existing owners to find spare parts or application programs for their existing goods.288

Network effects complicate legal analysis because, depending upon the characteristics of a given market and the strategies of the various firms, it may be difficult to tell whether a given policy is procompetitive. For example, there is evidence that computer game manufacturers accept low margins on their consoles in order to expand the market for compatible computer games.289 Depending upon the strength of copyrights, and the costs of switching to a different system, low-margin hardware pricing may allow manufacturers to earn more than they could by charging higher prices on their consoles. At the same time, expanding the ownership base of consoles might induce entry into the aftermarkets for software compatible with each manufacturer’s hardware. This would promote innovation in intellectual property and constrain monopoly pricing in the aftermarket. If either option is possible, what position should we take if consoles are priced below cost? In normal circumstances this might be sufficient to state a predation claim, but given the positive feedback effects with respect to providers of complementary goods (indirect network effects), can we say that a firm is likely to recoup its losses? And if the firm were likely to recoup its losses even in the face of entry and competition by firms providing complementary software of their own, thus producing a competitive aftermarket, is there really a problem for antitrust to cure?

The delightful economic puzzles presented by network effects take many forms, and this article cannot do them all justice.290 One aspect of network effects markets is particularly relevant to the essential facilities debate; this is the phenomenon of “tipping” among competing systems. As Professors Katz and Shapiro state the problem:

In markets with network effects, there is a natural tendency towards de facto standardization, which means everyone using the same system. Because of the strong positive-feedback elements, systems markets are especially prone to ‘tipping,’ which is the tendency of

288. Id. at 96.
289. Id. at 101.
290. These puzzles include the predation question mentioned in the text, the possibility of profitable leveraging strategies from ordinary markets into network effects markets, the possibility that network effects markets will be too small (suffer unexploited potential gains from trade) because existing users cannot compensate prospective users for the benefits they confer on existing users, the peculiar sensitivity of such markets to fluctuations in consumer expectations (all consumers want to own VHS rather than Beta machines) and thus the possibility of anticompetitive practices through manipulation of expectations, and the like.
one system to pull away from its rivals in popularity once it has gained an initial edge.\textsuperscript{291}

In and of itself, tipping is neither good nor bad from an economic perspective. If a firm owns the network to which a market tips, its increased revenues will offset the losses of firms whose products did not become the standard. As discussed above, this distributional consequence should be of no moment to antitrust.

Tipping does present the risk, however, that the system that becomes the standard will be or become inferior to some other system. If that happens, there is a risk of social welfare loss because consumers will be reluctant to depart from the standard for the simple reason that it is the standard, and consumers will face costs both in terms of old equipment abandoned and costs to retrain personnel in switching. Some thoughtful commentators have suggested that such forces are sufficiently strong that “computer operating systems serve as ‘essential facilities’ in computer hardware markets. Unless a firm can get onto the network, its products will be at a great disadvantage relative to those that can run the vast stock of application programs designed for the industry standard.”\textsuperscript{292}

Some believe there is empirical market evidence of such welfare losses, and the examples on which they rely are worth exploring in detail. In 1985 Paul A. David wrote an influential article suggesting that network effects could create social welfare losses by locking consumers into suboptimal technologies.\textsuperscript{293} His principal example, which quickly found its way into the legal literature,\textsuperscript{294} involves the configuration of keys on standard keyboards.\textsuperscript{295} It has been invoked by counsel for certain of Microsoft’s competitors in the course of arguing for enhanced antitrust scrutiny of that firm’s activities,\textsuperscript{296} and was cited by Judge Boudin in his First Circuit concurrence in *Lotus Development Corp. v. Borland International, Inc.*\textsuperscript{297} Keyboards exhibit network effects because the value of any given configuration depends

\textsuperscript{291} Katz & Shapiro, *Systems Competition*, supra note 250, at 106.

\textsuperscript{292} Menell, *Tailoring*, supra note 152, at 1366. Professor Menell’s later works suggest that he might have softened his position on this issue to some degree.

\textsuperscript{293} Paul David, *Clio and the Economics of QWERTY*, 75 Am. Econ. Rev. 332, May 1985 (Vol. 75, No. 2, Papers and Proceedings of the 97th Annual Meeting of the Amer. Econ. Assoc.).

\textsuperscript{294} E.g., Menell, *Tailoring*, supra note 152, at 1340.

\textsuperscript{295} Id. Known as the QWERTY configuration for reasons that will be obvious to the reader at a glance the next time a keyboard is encountered.


\textsuperscript{297} 49 F.3d 807, 820 (1st Cir. 1995), aff’d, 165 S. Ct. 804 (1996).
on how many people are trained to use it quickly. A firm that adopted a non-standard keyboard for its word processors would have to invest resources in training typists unfamiliar with the keyboard, and would have a harder time finding temporary typists when needed.

Insofar as it is relevant to software and antitrust, the argument is that the cost of switching from a dominant technology to an arguably superior technology may keep consumers locked into an inferior system, thus implying a need for an antitrust remedy. As one commentator has summarized it, the argument is as follows:

Primitive typewriters were unreliable mechanical devices, and the QWERTY keyboard was, therefore, deliberately designed to be dysfunctional so that typists would not strike the keys so rapidly that the device would jam. Obviously, modern software and computers can process keystrokes far more quickly, yet consumers are locked into the dysfunctional QWERTY standard. Superior keyboard layouts were developed years ago but were unsuccessful in dislodging the clearly inferior design that had established itself as an early standard.

The claim that “superior keyboard standards were developed years ago” rests upon a study by the U.S. Navy in 1943 that concluded a different keyboard, the Dvorak keyboard, was more efficient than the standard model. From this evidence, some have concluded that firms using the QWERTY keyboard have been incurring opportunity costs for 50 years. This is a strong contention. Why do keyboard consumers tolerate such costs? Could not schools and temporary agencies teach typing on both systems, in a way similar to the teaching of multiple word-processing programs? Are the costs of learning an additional keyboard pattern really so high? At least to the extent one believes that consumers are better judges of their utility functions than the Navy, there is reason to be suspicious: at a minimum we may safely conclude that the difference in efficiency between the two systems is not large; otherwise switching costs (which cumulate over time) would eventually be overcome.

There is more to it than that. There is good and perhaps compelling evidence that the entire QWERTY saga is a myth, the economic equivalent of alligators in the New York City sewer system. As Professors Liebowitz and Margolis tell it:

298. Reback, supra note 296, at 25.
299. Id.
[A]lmost every element of [the QWERTY] tale is false . . . . The QWERTY keyboard was not created to slow down typing speed. Early on, there were other publicized touch typists using other keyboards. The Navy study was very poorly documented and designed, and appears to have been conducted by Navy Lieutenant Commander August Dvorak, creator and patent holder on the keyboard bearing his name. A later, carefully constructed and controlled study, performed for the General Services Administration in the 1950s, demonstrated quite the opposite results from the Navy study. More recent studies indicate that there is practically no difference in typing speed between the two keyboard designs. The Dvorak typewriter keyboard, it turns out, is a rather poor empirical base upon which to support a theory.301

Anecdotes cannot supplant analysis, and the apparent disrobing of the QWERTY emperor no more establishes that antitrust may not be invoked to remedy supposed welfare losses in “tipped” markets than the anecdote itself could establish that antitrust should play such a role. Nevertheless, it is a cautionary tale that reminds us that consumer choices in the market deserve some respect. At a minimum, we must demand rigorous proof that the system consumers adopt is in fact inferior to a viable alternative system, and that the failure of that competing system to supplant the existing standard is in fact due to network effects. Without such evidence and analysis, strong conclusions about consumers being “locked into” a “dysfunctional standard” are not supportable.302 As yet, “the a priori case for network externalities is treacherous and the empirical case is yet to be presented.”303

Apart from the tenuous empirical foundation provided by QWERTY, however, the theoretical problem of tipping remains, although its extent and probability are still quite open to debate. The extremely low variable (and thus declining average) cost of producing software exacerbates the problem. The cost curve resembles that of natural monopoly, and thus presents a facially plausible case for

---

301. S.J. Liebowitz & Stephen E. Margolis, Network Externality: An Uncommon Tragedy, 8 J. ECON. PERSP. 133, 147 (1995) [hereinafter Network Externality]. Dvorak also conducted other studies, though his methodology suggested that he had particular outcomes in mind. In one study, Dvorak had students at the University of Chicago laboratory school trained on his keyboard and compared the speed at which they learned the keyboard with the speed at which students in conventional high schools learned the QWERTY keyboard. S.J. Liebowitz & Stephen Margolis, The Fable of the Keys, 33 J. L. & ECON. 1, 9 (1990).

302. See Liebowitz & Margolis, Network Externality, supra note 301, at 146.

303. Id. at 149.
invoking the essential facilities doctrine to compel access to the standard. For example, suppose one operating system dominated the market and that its architecture was closed—meaning that it asserted its copyrights against firms seeking to write application programs to work with the operating system. A firm that wrote such programs could contend that because of network effects, switching costs (which could be considered a collateral concomitant of network effects), and declining average costs, the operating system code was an essential facility to which the firm was entitled to access under the antitrust laws. Should such arguments prevail?

There are three reasons to reject the use of the essential facilities doctrine to compel access to software, even in the presence of network effects. The first, and relatively weak, reason is that software cost curves, though they resemble natural monopoly, do not reflect true natural monopoly conditions. Recall that a natural monopoly market is one in which a single firm can satisfy market demand at a lower cost than two or more firms. Firms in such markets do indeed tend to have declining average costs, but the reasoning may not be reversed; not all firms with declining average costs are natural monopolists. It follows that operating systems are replicable and thus do not fit the economic requirements for invocation of the essential facilities doctrine. The presence in the market of multiple operating systems is proof enough of this point. A firm that wishes to have its application software run on an operating system thus has the option of writing such a system. That is no doubt an expensive and time-consuming option. That potential entrants must replicate the capital expenditures of firms already in a market, and may find it difficult to do so, however, does not entitle them to subsidies from the existing firms.

Still, the natural monopoly argument is not as persuasive in the presence of network effects as it is in some other contexts (which is not to say that it is wrong). Because network effects may impede shifts

304. Empirical evidence does suggest, however, that the more tightly closed an architecture is the less likely it is to become a standard.

305. See supra notes 21 and 22.

306. This would be true even if the existing firm was allowed to charge a royalty for access to its code. The amount of the royalty would presumably have to be capped by reference to some benchmark. Otherwise the existing firm would simply demand a royalty high enough to keep the potential entrant out of the market. If it were profitable for the existing firm to charge such a royalty, it presumably would do so rather than close its architecture irrevocably. A firm forced to depart from a presumptively profit-maximizing policy of noncooperation for the benefit of its competitors thus may fairly be said to be granting them a subsidy, even if only in the form of foregone revenues.
from one competitive equilibrium (or standard) to another, entry into the market by new firms touting new potential standards will be less attractive (supply will be relatively less elastic) than would be the case in an ordinary market.\textsuperscript{307} To the extent that entry is deterred, the competitive process will work more sluggishly than would be the case without network effects. This means that a single firm possessing a standard in a network effects market might be extremely hard to dislodge, even though it would not be a natural monopolist. This line of argument helpfully focuses on the true risk to social welfare posed by the tipping phenomenon—the social loss comprised of the difference between the inferior existing standard and the superior new standard to which the market has difficulty shifting—rather than on the size or profitability of the firm owning the standard, which may be quite large.

Keeping this focus in mind, the question becomes whether antitrust can meaningfully improve the ability of the market transition from one standard to another when a superior standard appears. The first step in answering this question is to analyze how well transitions may be made without antitrust, which leads us back to \textit{Sega} and the ability of firms to reverse engineer software to achieve systems compatibility. If copyright allows firms to reverse engineer software in order to achieve compatibility, then new entrants will be able to build software bridges from old standards to new, facilitating transition and thus reducing (if not eliminating) the welfare loss associated with tipping. \textit{Lotus Development Corp. v. Borland International, Inc.},\textsuperscript{308} though different in important ways from \textit{Sega},\textsuperscript{309} is a case in point.

\textsuperscript{307} It should be stressed that network effects are by no means guaranteed to impede transitions from one efficient equilibrium to another. As Professors Katz and Shapiro note, “there is no general theoretical result implying excess inertia in market equilibria. Indeed, given the possibilities of multiple equilibria, markets may also exhibit the opposite of excess inertia . . . . [T]he market may be biased in favor of a new, superior, but incompatible technology.” Katz & Shapiro, \textit{Systems Competition}, supra note 250, at 108. The point is a theoretical one and the evidence in each relevant market must be examined to determine whether impediments exist.

\textsuperscript{308} 49 F.3d 807 (1st Cir. 1995), aff'd, 165 S. Ct. 804 (1996).

\textsuperscript{309} In particular, the interface code Accolade sought to incorporate in its computer games was itself purely functional and thus not copyrightable. Lotus's menu command hierarchy was arguably expressive and, prior to the First Circuit's ruling, copyrightable. And, unlike Accolade, Borland copied Lotus's software precisely for its expressive characteristics. Importantly, however, as Judge Boudin rightly noted in his concurrence, Borland did so only to permit Lotus users to transfer their own investment in learning Lotus's program to Borland's program. \textit{Id.} at 821 (Boudin, J. concurring). Similar to Accolade, Borland's software competed on its own merits, and thus did not deprive Lotus of a return on its investment in its software (as opposed to the investment of its customers in learning the software). \textit{Id.}
Much has been written about *Lotus*, and more is surely on its way; this is not the place to add to that debate. It is worth noting, as Judge Boudin did in his concurrence, that at one point Lotus's 1-2-3 spreadsheet software "had such sway in the market that it has represented the *de facto* standard for electronic spreadsheet commands." Borland's copying of Lotus's menu command hierarchy allowed consumers to transfer their investment in learning Lotus's system to Borland's spreadsheet, and thus avoid the cost of re-learning material they already had spent the time and money necessary to learn. The law in this area is obviously in flux, and any legal equilibrium built upon a 4-4 tie in the Supreme Court is quite obviously unstable. But so long as copyright countenances limited copying for the purposes of achieving hardware compatibility, or building bridges between software standards, software markets will find it easier to transfer from old equilibria to new. Social welfare losses from network effects will thus be minimized, significantly diminishing the force of any argument for bringing antitrust principles into the picture.

The final reason to reject market tipping as a basis for invoking the essential facilities doctrine in software network effects markets is that such invocation could well make the network effects even stronger, and thus make entry (and thus market discipline) more sluggish. Antitrust cannot determine whether an existing standard is

---

310. It is worth noting, however, that *Lotus* adopted the relatively aggressive holding that Lotus's menu command hierarchy was "an unprotectable method of operation" under section 102(b) of the Copyright Act. *Id.* at 818-19. This holding obviously facilitates compatibility, and will thus aid in innovation and reduce the lock-in power of network effects. At the same time, there is some risk that this will induce (or at least condone) behavior inconsistent with copyright; Borland (like Accolade) was competing with Lotus on the merits of Borland's own spreadsheet software, which was itself protected under copyright. Allowing Borland to build a bridge between its program and Lotus's facilitated innovation by making it easier for Lotus users to transfer their investment in learning a system to Borland's program (arguably superior, based on market evidence). The situation arguably should be different if Borland did not attempt to compete on the merits of its own creative work, but instead simply cloned Lotus's program and sold it in a Borland package. *See id.* at 821 (Boudin, J., concurring). Fair use, which is a flexible concept that examines numerous variables, would provide a more tailored approach to the problem, capable of distinguishing between cases of bridge-building and cloning. Precisely because of its flexibility, however, fair use outcomes are more difficult to predict than the court's relatively bright-line holding under section 102(b). With uncertainty comes risk, and the possibility that some firms would avoid normatively legitimate bridge-building for fear that a court would wrongly perceive its actions to be illegitimate cloning and thus reject a fair use defense. *Id.* In the event other circuits choose fair use as a more appropriate approach to the problem, the market preemption analysis discussed above in connection with *Sega* would apply to such cases.

311. *Id.* (emphasis in original).
efficient or should be abandoned in favor of a superior newcomer. At
most, the essential facilities doctrine could compel a firm owning a
standard to grant access to other firms on some economically
reasonable basis. If the owner of an operating system were precluded
from closing its architecture to firms seeking to write complementary
programs, more programs would likely be written for the operating
system, reinforcing one of the factors presumed to cause tipping. In
other words, it makes no sense to attempt to remedy inertia by adding
to the load of goods that tie the market to the existing system in the
first place.

IV
Conclusion
“Don’t just do something, stand there!”

Dean Acheson

The rapid evolution of computer software, its increasing
importance in the economy as a whole, and its novelty relative to
traditional forms of copyrighted works, contribute to a potentially
destructive atmosphere of urgency and confusion in the debate over
how the law should deal with competing software firms. To add to the
difficulty, the monetary stakes in these cases are enormous, including
the risk that an erroneous decision pushing protection for software too
far to either end of the spectrum will hamstring innovation and turn
what could be a smooth and efficient network into a Tower of Babel.
Add in the ambiguous economics of such markets, including the very
real possibility that positive feedback effects could cause any given
decision to boomerang on a court, and one is left with the conclusion
that the courts deserve very high marks for doing as well as they have.
Sega and Lotus promise to relieve many of the strains that could arise
in these markets; the ISO cases have properly rejected the essential
facilities concept, and the affirmation of copyrights in diagnostic
software, while unfortunate from a policy perspective, is essentially
compelled by the existing statutory scheme.

312. Thoughts on the Business Life, FORBES, Dec. 1., 1977, at 140 (attributing quotation to
Dean Acheson).
There is no room for sanguinity, however. The Court was only one vote away from what might have been an entirely different outcome in *Lotus*; and Justice Stevens, who recused himself from the case, has some history of disfavoring bright-line rules of the type adopted by the First Circuit. Apart from what may have been a narrow escape for the cause of compatibility, the complex and confusing aspects of competition in software markets tends to create a desire on the part of both lawyers and judges for familiar and seemingly established rules to apply. In such an environment, the law's tendency to borrow concepts from one field to fill a gap in another is heightened. This borrowing is always risky, but is made riskier by balkanization within academe. Experts in a given subject, such as copyright, have an understandable tendency to focus on those features of the subject that need mending. Focusing on (and writing about) what is right about a given doctrine is both less challenging and less useful in pursuit of tenure. Unfortunately, this may lead to a belief that the situation in the expert's field is so bad that things simply have to be better elsewhere; thus copyright experts sometimes look to antitrust to provide remedies for problems they fear will be exacerbated by copyright, and the uncritical migration of legal concepts is thus given academic imprimatur.

At the risk of self-parody, I believe that under the present state of the evidence and the present economic understanding of how network effects markets transition between equilibrium points, the essential facilities doctrine has no place in the legal regime being crafted to regulate software. There are admittedly instances in which intellectual property rights have been asserted to defeat competition in circumstances where the purposes of the intellectual property laws are not served. And the evolving economic understanding of network effects does suggest the possibility of social welfare losses in software markets. But there is as yet no basis to believe that antitrust can avoid such losses. In addition, the fundamental issues raised by these questions involve the proper scope of intellectual property rights.

313. His dissents in *Brooke Group, Ltd. v. Brown & Williamson Tobacco Corp.*, 509 U.S. 209, 243 (1993), and *City of Columbia v. Omni Outdoor Advertising, Inc.*, 499 U.S. 365 (1991), both of which reversed jury verdicts for antitrust plaintiffs, tend to support this thesis. On the other hand, Justice Stevens is also the author of *Sony Corp. of America v. Universal City Studios, Inc.*, 464 U.S. 417 (1984), which established a strong presumption (since weakened by *Acuff-Rose*) that commercial uses of copyrighted material were not protected by fair use.

314. See Menell, *Tailoring, supra* note 152, at 1366 (suggesting application of essential facilities doctrine to operating systems software).
Antitrust simply does not have the analytical tools necessary to provide coherent answers, which can only be derived from the principles and goals of the intellectual property laws using the analytical approach those laws embody.