Does Antitrust Need to Be Modernized?
Author(s): Dennis W. Carlton
Source: The Journal of Economic Perspectives, Vol. 21, No. 3 (Summer, 2007), pp. 155-176
Published by: American Economic Association
Stable URL: http://www.jstor.org/stable/30033739
Accessed: 06/10/2011 13:42

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

American Economic Association is collaborating with JSTOR to digitize, preserve and extend access to The Journal of Economic Perspectives.
Does Antitrust Need to be Modernized?

Dennis W. Carlton

Competition is at the heart of the U.S. economic system and the antitrust laws influence how that competition takes place. The Clayton Act forbids mergers that are anticompetitive. The Sherman Act forbids the formation of cartels. It also forbids certain conduct that a firm can use to maintain or create market power, though it remains perfectly legal for a monopoly to charge a high price if that firm achieved its monopoly fair and square. The Federal Trade Commission (FTC) Act forbids unfair methods of competition. Within the last 30 years, the courts’ interpretation of the antitrust laws has been heavily influenced by economic reasoning, especially from the field of industrial organization.

Antitrust doctrines raise fundamental questions about what economists know and do not know about competition. Should defining “markets” and calculating market shares remain a pillar of antitrust policy? Do economists understand enough about the effect of industry concentration on pricing or on R&D that numerical guidelines based on market concentration make sense? Do economists understand enough about abuses of the patent system so as to adjust antitrust policy related to intellectual property? Do economists know enough about some particular business practices such as tying—one focus of the recent Microsoft antitrust case—that some practices should be outlawed? How can economists fashion remedies for antitrust problems to increase their confidence that the remedies do not cause more harm than good?

This paper seeks answers to these difficult questions by considering controver-

---

Dennis W. Carlton is Professor of Economics, Graduate School of Business, University of Chicago, Chicago, Illinois and Research Associate, National Bureau of Economic Research, Cambridge, Massachusetts. He is currently serving as Deputy Assistant Attorney General for Economic Analysis, Antitrust Division, U.S. Department of Justice, Washington, D.C. His e-mail address is (dennis.carlton@ChicagoGSB.edu).
sial antitrust doctrines that need fixing, or at least some modernizing. Specifically, I analyze market definition; the interaction of intellectual property and antitrust law; certain types of exclusionary conduct (tying and bundling discounts); and procedural issues involving economic matters such as damage multiples, the right to sue, and laws of contribution.

My opinions in this paper have been shaped not only by my academic and consulting experience, but also by my experience as the Deputy Assistant Attorney General for Economic Analysis and as a Commissioner on the Antitrust Modernization Commission (AMC), which Congress established in 2002 to investigate whether the antitrust laws and their administration need to be modernized. Congress was particularly interested in whether rapid technological change and globalization require new laws or approaches to antitrust. The AMC is comprised of twelve members, eleven of whom are, or recently have been, practicing lawyers. I was the sole economist. The Commission issued a report in April 2007 based on extensive hearings, available at (http://www.amc.gov/). The views expressed in this paper are my own and do not necessarily reflect those of the AMC or those of the Antitrust Division of the Department of Justice.

What Is the Objective for Antitrust and How Should It Be Achieved?

Before discussing a selection of what I believe to be the most interesting antitrust topics, I begin with two policy questions: First, what should be the objective of the antitrust laws? Second, why is antitrust law organized around certain “safe harbor” behaviors that are almost always allowed and certain actions that are “per se” illegal, with a zone of discretion for antitrust enforcement agencies and courts?

The Objective: Total Surplus

Antitrust laws influence how firms compete. For example, the Robinson–Patman Act was passed in 1936 in response to small firms’ complaints about the ability of large firms such as the grocery chain A&P to obtain low prices from suppliers. By inhibiting large firms from obtaining discounts unless they could be “cost justified,” the Robinson–Patman Act has harmed consumers through higher prices, but protected small firms.\(^1\) As another example, consider a policy that encourages mergers that can lead to cost savings, which in turn lead to lower prices for all consumers. Clearly all consumers benefit. But in many mergers, some consumers may be harmed while others benefit. Consider a proposed airline merger that will lead to a very efficient route structure but will also result in less service to some cities. Even if most passengers are benefited, some are hurt, so the

\(^1\) The reader might surmise that repeal of the Robinson–Patman Act would be desirable, a position on which economic analysis seems virtually unanimous (Posner, 2001).
question arises whether some consumers should be weighted more heavily than others in deciding whether to allow the merger to occur.

Should antitrust seek to maximize consumer surplus, total surplus, or some weighted average of producer plus consumer surplus? The Department of Justice and the Federal Trade Commission, the two federal antitrust agencies, often state that their focus is on consumers, which seems to imply a focus on consumer surplus. My experience is that although technically, under the Clayton Act, antitrust harm to any substantial consumer group could provide a basis to block a merger, the antitrust agencies mainly look at aggregate effects. In other words, if consumers gain as a group, antitrust agencies generally do not distinguish amongst consumers.

Antitrust agencies also suggest that they might not challenge activities—such as mergers—that promise unusually large efficiencies, even if some consumers are harmed (U.S. Department of Justice and FTC Horizontal Merger Guidelines, Section 4, issued 1992, revised 1997). However, U.S. courts generally have not recognized efficiencies as a defense to antitrust activity that harms consumers.³

The proper objective of antitrust should be total surplus, not consumer surplus (Heyer, 2006). The fundamental reason is familiar to most economists: it is better to pursue public policies that maximize output and then worry about distributional questions, rather than to pursue inefficient policies. I will first lay out some of the arguments against a pure consumer surplus standard and then address some of the counterarguments in favor of such a standard.

The first and perhaps the most significant practical problem with a consumer surplus standard is that, as commonly applied, it tends to favor short-run price reductions over long-run efficiency gains. For example, it is commonly believed that when government antitrust authorities assess a potential merger, they focus on price effects over a two-year future period. Suppose that the merger offers two kinds of gains: a saving in fixed costs and a saving in marginal costs. Under a consumer surplus standard, only the saving in marginal costs will carry weight because it will reduce prices, while the fixed-cost savings is not considered as a benefit to consumers. But many high-tech industries have high fixed costs and low marginal costs—and although they develop new products rapidly, their new product cycle is often more than two years. Gains that lead to lower fixed costs today can encourage research and development, new products, and plants in the future. However, by focusing only on efficiencies that influence price over a short period, a government antitrust agency risks failing to credit the future efficiencies that will benefit consumers in the long run. To put it another way, the fixed-cost savings of today are the variable-cost savings in the future for new products.

³ There is semantic confusion in the economic and legal literature with some writers such as Posner (2001) and Bork (1978) using the term “consumer welfare” to mean total surplus, while others use it to mean only consumer surplus.

³ The antitrust laws of most countries focus on consumer surplus, rather than total surplus. Canada and New Zealand are the rare exceptions in that they use a total surplus standard. Canada has experienced considerable litigation over the meaning of total surplus, with the outcome finally reached that the Canadian Competition Bureau can use total surplus in the sense that economists do.
Of course, it would be theoretically possible to take a long-run perspective on future consumer surplus that would include future gains from new technologies and products, and then to estimate the present discounted value of such gains. But such a calculation will depend on the difficult-to-estimate benefits of future products. Focusing on total surplus, even if only in the short run, will better encourage government antitrust agencies to recognize fixed-cost savings as a source of future benefits to consumers.

A second argument against a consumer surplus standard is that thinking of antitrust as protecting innocent individuals from evil corporate empires is misleading (though sometimes effective as rhetoric). Most transactions in the U.S. economy are between firms. Firms are typically both the consumers and the sellers. Moreover, firms are owned by shareholders, so profits do flow back to households. The use of total welfare treats all agents in the economy the same, showing preference to no particular group. The use of consumer surplus shows preference to consumers over producers.

A final argument against a consumer surplus standard is that, if only consumers matter, then a buying cartel should be perfectly legal and indeed should be encouraged. Monopsony power would not matter in antitrust cases, because the fact that sellers are harmed is irrelevant under a consumer surplus standard. I know of no proponent of the consumer surplus standard who endorses buyer cartels, or who believes that monopsony is not harmful. Instead, proponents of a consumer surplus rule tend to argue that buyer cartels and monopsony are exceptions to the otherwise sensible rule of maximizing consumer surplus. However, the need for these exceptions illustrates the lack of a coherent logic for the consumer surplus standard.

There are several counterarguments in favor of choosing a consumer surplus standard: that it doesn’t matter much; that it is easier to monitor antitrust authorities who focus on consumer benefits; and that a focus on consumer surplus is a political necessity. I consider these arguments in turn.

As a practical matter, how much difference does it make if one focuses on consumer surplus, not total surplus? For most situations, both standards will lead to a similar result. After all, many actions that achieve efficiencies for firms should be expected to help consumers (for the evidence on merger efficiencies, see Carlton and Perloff, 2005, chap. 2). Even in those cases where an activity like a merger would pass the total surplus standard but not the consumer surplus standard, the firm engaging in the action has enough resources to pay the consumers to make them better off. Indeed, some merging firms now undoubtedly go to their major customers and, by offering desirable long-term pricing, eliminate the customers’ opposition to the merger.

However, it would be unwise to be too sanguine about how bargaining between firms and customers will lead to efficient antitrust outcomes. After all, a variety of bargaining games are possible. For example, customers might assess whether their complaint to a government agency could scuttle an entire merger and, if so, demand the total surplus from the deal. In fact, customers who fail to coordinate
their demands may collectively demand more than the total surplus from the deal. Conversely, if buyers anticipate that their opposition will not scuttle the deal, then they may accept a pittance not to complain. If the government agencies rely only on the lack of customer complaints in deciding whether to approve a merger, then in this case, deals that harm welfare can be approved. What this analysis does indicate, though, is that the possibility of bargaining means that the number of cases where it matters whether one uses a consumer surplus or a total surplus standard may be even smaller than it first appears. It also illustrates that a government agency must examine why customers are (or are not) complaining (Heyer, 2006).

The most potent reason to support a short-run consumer surplus standard relates to the monitoring of antitrust policy. If an antitrust agency adopts a short-run consumer surplus standard, then it is possible to monitor the agency to some extent by seeing whether consumers are harmed in the short run by elevated prices. If instead one adopts a short-run total surplus standard, it will be more difficult to verify whether agency officials are achieving their objectives. A great deal of information about the effects on consumers can be gained by looking at changes in market prices (though the need to hold other factors constant can make this analysis trickier than it may sound). In comparison, the measurement of efficiency gains to firms, as required in a total surplus standard, is harder to verify. The difficulty with a long-run consumer surplus standard is that by the time one has determined whether long-run surplus has increased, the officials responsible may have moved out of the antitrust agency so that there is no one to discipline. Therefore, in countries where judges or government agencies will be susceptible to political influence, a short-run consumer surplus standard might lead to higher total welfare than a total surplus or long-run consumer surplus standard because it is easier to monitor decision makers with the short-run consumer surplus standard.

A populist objection to total surplus as an antitrust objective is that it is less favorable to consumers than a consumer surplus standard. As discussed earlier, I believe this populist justification is based on false premises. A short-run total welfare standard is more likely to maximize long-run consumer surplus than is a short-run consumer surplus standard. This outcome is especially likely in a dynamic economy where new products are the primary way that consumers benefit.4

Safe Harbors and “Per se” Rules in an Environment with Costs and Uncertainty

The legal system involves costs, both out-of-pocket costs and costs of making errors. Moreover, the cost of errors must include not only the cost of mistakes on

4 A clever theoretical insight from Lyons (2003) is that firms choose which mergers to pursue subject to antitrust constraints. The profit-maximizing merger in the feasible set can differ depending upon whether consumer surplus or total surplus is used as an antitrust criterion. The empirical significance of this point, and whether it suggests that consumer or total surplus is the better criterion, is ambiguous. See also Farrell and Katz (2006).
the firms involved in a particular case, but also the effect of setting a legal precedent that will cause other firms to adjust their behavior inefficiently.

The recognition that a legal process has costs and can commit error implies that we would not want courts to engage in a detailed investigation of every pricing or marketing decision of a firm. For example, imagine that every decision of a firm to reduce prices could be challenged as potentially anticompetitive “predatory pricing.” Firms might decide to minimize all price-cutting behavior out of a fear that a court might find them guilty of predatory pricing. This fear could chill price competition among firms. Hence, antitrust authorities and courts do not investigate firms for price cutting, as long as price is above “cost.” (Let’s set aside the question of what measure of cost should be used.) Even though one can easily construct theoretical models of above-cost predatory pricing, antitrust authorities treat above-cost pricing decisions as a safe harbor, not to be challenged.

A similar logic applies to entry decisions. It is theoretically possible that if an inefficient firm enters a monopoly market, total surplus may fall (that is, the inefficiency of the new firm may more than offset any gains from greater competition). However, entry is so vital to competition that subjecting firms to possible legal liability for entry is unwise policy. The potential loss from chilling this form of competition far outweighs any benefit from those few cases where entry does harm efficiency—even assuming the court can accurately identify those cases.

Figuring out what should be safe harbors for competitive behavior depends on judgments about how error-prone courts are, how costly litigation is, and how vital the attacked behavior is to competition. My own view is that markets are generally better than courts at producing competition (see also Easterbrook, 1984), and therefore, for certain acts such as entry, pricing, and product innovation, safe harbors generally make lots of sense. This point holds even though numerous academic articles, including my own, show the theoretical possibility of social harm from strategic use of these actions in certain circumstances. On the other side, just as there is a rationale for safe harbors to protect actions that are unlikely to harm competition, so too there is a rationale for “per se” rules to forbid actions that are almost always anticompetitive, such as explicit price fixing.

Thus, in thinking about how to achieve the antitrust goal of maximizing total surplus, in a number of cases, setting up safe harbor rules for permitted actions and per se rules for prohibited actions will be more sensible than attempting to do a full analysis of every business decision. Let me now turn to a discussion of some of the controversial topics in antitrust.

What Role Should Definition of Markets Play in Antitrust?

Courts often analyze antitrust cases by determining the relevant market and then calculating shares of different firms within that market. This calculation remains central in the legal process to evaluating whether a firm or group of firms has market power—the ability of a firm or group of firms acting together to raise price profitably
above competitive levels. For example, some courts use market shares as a screen at the
time of summary judgment to decide whether to allow a case to go forward.

The definition of a market can determine the outcome of an antitrust case. The classic "cellophane" case offers a vivid illustration. The U.S. government charged that du Pont monopolized interstate commerce in cellophane in violation of Section 2 of the Sherman Act. The government showed that du Pont produced almost 75 percent of the cellophane sold in the United States. Du Pont's defense was that even if it had a dominant position in cellophane, the relevant market should consider all flexible packaging materials, and by that standard, cellophane constituted less than 20 percent of all flexible packaging materials sold in the United States. The U.S. district court accepted this argument and dismissed the case, and the U.S. Supreme Court affirmed that dismissal in United States v. E.I. du Pont de Nemours Co. (351 U.S. 377 [1956]). The Court failed to recognize that there are at least two possible questions: "Does du Pont have the ability to raise price profitably above the current price?" and "Is du Pont setting the current price above competitive levels?" The Court actually answered the first question, even though it thought it was answering the second one (Carlton, 2007). This error is known as the "cellophane fallacy."

As another example of the importance of market definition, consider the case United States v. General Dynamics (415 U.S. 486 [1974]) in which two producers of coal sought to merge. Coal is often sold pursuant to long-term contracts. The Court concluded that a high market share based on a market defined as coal produced was an incorrect indicator of a firm's competitive significance, and that the correct indicator was a firm's share of uncommitted, not yet contracted, coal reserves. In other words, a coal producer who has already committed to sell all its coal at a fixed price is of no competitive significance in establishing future prices.

A loose economic definition of a market is that it comprises all those products whose presence constrains the price of a particular product to a particular level. For economists, drawing bright line boundaries around products in a market often makes no sense. Indeed, if antitrust law did not commonly require defining a market, economists would probably spend much less time discussing what the denominator of a market share should include. Instead, economists would try to estimate demand systems econometrically to get a sense of substitution patterns amongst different products and then use that knowledge to estimate the effect of a merger or some other questioned business practice. In comparison with this kind of analysis, market shares are at best a crude first step.

The crude nature of market shares as a tool to analyze market power is well understood by government agencies and some courts. However, some courts are likely to be less sophisticated than the government agencies in evaluating detailed econometric studies of whether a certain merger or business practice is anticompetitive. In such a setting, using crude market share analysis as a screen for deciding whether to allow cases to go forward may be sensible for a court.

There are three separate circumstances where the use of market definition merits discussion: horizontal mergers; strategic behavior by single firms; and new technologies. Only in the first circumstance is market definition immune from
serious flaws, though even there, problems can arise. In the other cases, the ability of economists to define a market is quite limited (Carlton, 2007).

**Horizontal Mergers**

The Merger Guidelines of the U.S. Department of Justice and the Federal Trade Commission go through an elaborate and, for the most part, well-reasoned method for defining a market. Basically, a market has the property that, absent entry, a monopolist of the product or products in the market would raise price by a significant amount (for example, 5 percent) above current levels for a significant period (for example, two years). Once the market is defined, the next step is to calculate market shares. If two firms with sufficiently large market shares merge, then there is a presumption that prices will rise from current levels (although this presumption is rebuttable, as discussed in the Merger Guidelines at §1.51). Notice that the benchmark is the current price, not some (unobserved) competitive price.

The issue of market definition arises in virtually every merger case. In the recent merger of Whirlpool and Maytag, the question arose as to whether front-loading washing machines are in the same market as top-loading washing machines. In the recent merger of SBC and ATT, the question arose as to whether cell phones are in the same market as landline phones. In mergers of movie theatres, the question arises as to how close together two movie theatres need to be in order to be considered to be in the same market.\(^5\)

The Merger Guidelines do a good job of defining the properties of a market in a merger case. Whether a methodology can be devised to construct a market with such properties is another matter. The definition would seem to require an econometric estimation of a demand system for related but perhaps differentiated products. Then, these demand estimates can be used to figure out what set of products have the property that a monopolist of that set of products could profitably raise price by 5 percent.\(^6\) Given a market definition, one then calculates market shares to make inferences about the effect on prices of a proposed merger. Upon reflection, this series of calculations may seem rather odd. After all, this approach requires proceeding through a complicated econometric demand analysis—and then basing predictions about the effect of a merger on a general intuition based on market shares.\(^7\)

---

\(^5\) I served as an expert for the firms in these recent matters.

\(^6\) Suppose that a market includes several possible products. Is the subset of possible products that comprise a market unique? No. Therefore, one must add some additional criterion such as minimum number of products or smallest value of commerce to obtain uniqueness. If products are not homogenous, should each product’s price rise by 5 percent or should some index rise by 5 percent? This question is not answered by the Merger Guidelines.

\(^7\) Another approach to market definition is to ask consumers to which products they would substitute if the price of the product under analysis were to rise 5 percent. One can then consider all those mentioned products as being in the market. Although this approach may be easy to implement, it is not equivalent to that in the Guidelines because the Guidelines include products in the market only if substitution to those products is sufficient to make unprofitable, say, a 5 percent price increase of the product under analysis, while the alternative approach ignores the strength of the constraining effect on price of the products to which consumers switch. The fact that this approach is not equivalent to the Merger Guidelines’ definition does not seem to be well understood.
But once the demand system is estimated, it can be used in conjunction with assumptions, including those about the competitive game, to yield structurally-based estimates of the effect of a proposed merger. This approach of merger simulation is conceptually well-founded on the underlying economic structure, though it has limitations because of the many assumptions on which it is based (Carlton, 2003, 2004). Still this approach can be viewed as a way to understand the pricing constraints imposed by the demand system. Basing predictions about the effect of a merger instead on a general intuition based on market shares is much cruder.

Sometimes a better way of proceeding is to use the market definition to calculate market shares and then relate those shares econometrically to price, using either a time series or cross-section. The idea is to see whether differences in market structure over time or across different locations lead to different pricing. This traditional price-concentration study looks a bit like a reduced form analysis. The main econometric issue is whether the market share can be treated as an exogenous variable affecting prices. If it is not exogenous, then one must find instrumental variables. For example, in a proposed merger of two railroads that will reduce the number of railroads from three to two, one might be able to compare the rates on railroad routes where there are currently three railroads to rates on other routes where there are only two railroads (adjusting for other factors) in an attempt to assess the effect of the merger. Endogeneity is not likely to be an issue, since the decision to lay track was typically made many years ago. When the number of competitors is endogenous, instruments reflecting regulatory barriers or entry costs can sometimes be used successfully. But as long as the endogeneity of market share can be addressed, this approach can sometimes be a reasonable way to estimate the short-run effects of a merger. The effects predicted from such an analysis can be checked for consistency with results from the econometric demand estimations and merger simulation.

**Strategic Behavior by Single Firms**

The antitrust laws in certain circumstances may deem as anticompetitive a wide range of single-firm strategic conduct, such as exclusive dealing, exclusive territories, nonlinear pricing, tying, and predatory pricing. Such cases are usually brought under the Sherman Act and are referred to as “bad act” cases. There are usually two requirements for a violation: 1) substantial market power, and 2) that the “bad act” maintains or enhances the market power.

Three cases illustrate “bad act” cases: In *U.S. v. Dentsply* (399 F.3d 181 [3d Cir. 2005]), a manufacturer that sold over 75 percent of all artificial teeth, had exclusive dealing arrangements with distributors of false teeth. The court ruled that this exclusivity prevented other manufacturers from competing effectively. In *Berkey Photo Inc. v. Eastman Kodak* (603 F.2d 263 [2d Cir. 1979]), Berkey complained that Kodak violated the antitrust laws by failing to inform Berkey of the characteristics of its new film format prior to its introduction, thereby making it difficult for Berkey to compete with Kodak in the developing of the new film. The court ruled
against Berkey. Finally, in *U.S. v. Microsoft* (253 F.3d 34 [D.C. Cir. 2001]), the government alleged that Microsoft misused its market power in its Windows operating systems through a variety of tactics including tying its browser (Internet Explorer) to maintain its market power. Microsoft was found liable (Whinston, 2001).

Often, analysts try to adapt the Merger Guidelines to the problem of defining markets in a “bad act” setting. However, the definition of markets in the merger setting refers to whether a monopolist would raise the price by, say, 5 percent above its current level. In contrast, in looking at “bad acts,” using the current market price as a benchmark will not work if the existing price already reflects market power achieved through the “bad acts.” Thus, some analysts have tried to adapt the definition of a market in the “bad acts” context to include all those products such that a hypothetical monopolist of the products could profitably raise price 5 percent above the competitive price. But this approach is an analytical dead end. If one knows the competitive price, there is no need to implement the definition of a market, since it would be obvious whether current price exceeds the competitive price. On the other hand, if one does not know the competitive price, this definition cannot be implemented!

A further difficulty is that most “bad act” cases involve monopolistically competitive firms selling differentiated products. Since monopolistically competitive firms have some market power in the sense that price exceeds marginal cost, presumably the deviation between price and marginal cost (the size of the dead-weight loss would be a superior measure) should be significant if it is to expose the firm to antitrust scrutiny. But no consensus exists in the courts or among economists as to how large this deviation should be. Maybe the courts should focus on whether profits are excessive? Yet for courts to attempt the difficult calculation of economic rates of return strikes me as not generally helpful. Arguments about the relevant time frame and accounting issues would make such analysis extremely difficult.

The difficulty of detecting market power in these “bad act” cases is even more complicated than I have described, because the alleged “bad act” may improve product quality (or promotion)—and appear in the short run to raise price. For example, a standard justification for exclusive territories is to provide incentives to the exclusive distributors to advertise and otherwise promote effectively. Thus, relying on price comparisons with and without the alleged “bad act” fails to address the relevant economics, because the “quality” (or marketing) of the good may not be held constant. It’s possible to design sophisticated statistical tests to circumvent this problem by modeling demand as a function of both price and quality, but such tests also raise the costs of the process and the risk of error.

Courts often rely heavily on market definition as a screen as to whether to proceed to a more sophisticated analysis of “bad acts.” It might sometimes be better to alter the sequence of analysis and first ask whether the conduct should be immune from antitrust challenges (even if there is market power); then if it should not be immune, ask whether there is significant market power; and if so, perform
a full analysis of the “bad act.” For example, in the Berkey case, the court ruled that Kodak was under no obligation to inform its rivals about future product introductions because such an obligation would lower Kodak’s incentive to innovate.

Antitrust, Technology, and Research and Development

Many mergers, like those in the telecommunications industry and several recent cases such as the Microsoft case, are in industries where technological change is key. The problem here is that when goods in the market are changing rapidly, defining the market in a defensible way becomes especially difficult.

One approach is to define an “innovation market” consisting of resources devoted to research and development in the relevant industry. The U.S. Department of Justice followed this approach when ZF attempted to acquire the Allison Transmission Division of General Motors, for whom I consulted. Both ZF and Allison made automatic transmissions for certain types of trucks. The Department of Justice claimed that combining the two firms would adversely affect competition for innovation in automatic transmissions (Complaint, United States v. General Motors, No. 93-530 [D. Del. 1993]). The deal was abandoned. This approach switches the market definition from products to inputs. This definition might be sensible if there were a clear connection between the input known as research and development and the eventual output. But this linkage is typically not clear, except perhaps in a few industries such as pharmaceuticals where the phases of drug development and approval often follow a timeline determined in part by government regulations. Quite often, innovations come from outside the specific industry under discussion. Attempts to link concentration of research and development in one particular area to speed of discovery are on much less solid footing than the (already weak) empirical base linking concentration to undesirably high pricing (Gilbert, 2005). Accordingly, the use of innovation markets as a way to measure market power in industries undergoing rapid technological change has little theoretical or empirical support (Carlton and Gertner, 2003).

Does Antitrust Doctrine Need Adjustments for Intellectual Property?

Property rights in intellectual property are designed to create incentives to innovate. However, the premise that greater protection of intellectual property necessarily fosters more innovation turns out to be false. Some protection of intellectual property is necessary to encourage innovation, but too much protection can inhibit innovation. For example, if obvious ideas are patented, then subsequent innovations that rely on these ideas will be forced to pay for the use of these obvious ideas, which in turn reduces the incentives for innovations that build upon these ideas. There has been much criticism of our current patent system, especially with regard to the lax application of a nonobviousness standard, which, it is claimed,

The question for this paper is whether antitrust strikes the right balance between encouraging the sort of market power that provides an incentive for innovation and discouraging the sort of market power that inhibits innovation. There are several ways in which firms can behave strategically to take advantage of our flawed patent laws, and modernized antitrust may have a role to play here. I discuss three possible areas for improvement: settlements, standard setting, and remedies. There may be no more important topic in antitrust than its relation to intellectual property, especially given how economic growth depends on innovation.

Settlements

Consider two types of circumstances involving settlements of patent suits that appear quite similar to an outsider, although one is anticompetitive and one is procompetitive. Suppose Firm A has a patent on a product. Firm B starts producing that product and claims that Firm A’s patent is not valid. Firm A sues Firm B. However, suppose Firm A—suspecting that its patent will be invalid—says to Firm B, “Listen, let’s settle the lawsuit. Why don’t you stay out of the market for this product, let me reap monopoly profits, and I will give you some of the profits.” This “settlement” of the lawsuit involves Firm A paying Firm B, the alleged infringer, to cease to be a competitive force in the market (Lemley and Shapiro, 2005). Moreover, if part of the settlement requires Firm B to make royalty payments to Firm A before withdrawing from the market, and then a subsequent Firm C produces the product, then Firm A can use any royalty received from Firm B as evidence of the patent’s validity.

This anticompetitive set of circumstances contrasts with an alternative set of circumstances. Suppose that Firm A believes that its patent is likely to be judged valid and infringed, but the risks and costs of lingering litigation are so high that the Firm A decides to pay Firm B an amount less than the projected cost of litigation. Getting rid of this nuisance lawsuit may well be in the patentee’s (and society’s) interest under certain circumstances.

In *Schering-Plough v. FTC* (402 F.3d, 1056 [11th Cir. 2005]), the Federal Trade Commission alleged that Schering-Plough violated the antitrust laws because it settled a patent dispute against two claimed patent infringers by, in part, paying them to delay their entry, thereby harming competition. The Court of Appeals ruled against the Federal Trade Commission and noted that patent settlements can serve valid business purposes.

As the *Schering-Plough* case illustrates, distinguishing between the pro- and anticompetitive effect of a patent settlement can be difficult, but some progress is possible. To prevent consumers from being worse off from patent settlements, the law could require that settlement agreements in which Firm B never enters the market in return for a payment are per se illegal, but agreements in which Firm B receives no payment but is allowed to enter before patent expiration, are acceptable. (Presumably, the stronger is Firm B’s position, the earlier its allowed entry
time will be.) The latter type of settlement guarantees that the only way the alleged infringer can benefit from the settlement is if it enters and competes to the benefit of consumers. This answer is not complete, but it would mitigate the use of patent settlements as a cartelization device.

**Standard Setting**

Some specialized areas now have so many patents that firms are in constant danger of facing a patent infringement claim from their competitors. A common reaction is for firms to develop their own patent portfolios, which they use to bargain with other firms over patent protection. Such “patent thicket” problems will discourage new firms that do not have a portfolio of patents to trade. These problems are a result of the patent laws, not antitrust (Shapiro, 2001); however, antitrust can make a difference even here. This subsection discusses standard setting; the next subsection discusses remedies when a patent is found to have been infringed.

Some standard-setting organizations wish to set industry standards that will not trigger large royalty payments to patent holders. However, so many patents are filed or soon to be filed that it is not easy to be sure that a particular standard does not (or will not) trigger such payments. In such situations, some firms have expressed a desire to agree in advance amongst themselves either not to charge a royalty or to charge at most only a small specified royalty should one of their patents be triggered by the standard. Firms have expressed concerns about antitrust liability for such arrangements. Courts should generally encourage such agreements and recognize their efficiency, while preserving the ability to condemn an arrangement that is a sham price-fixing agreement. The Department of Justice recently considered such a proposal by a standards development organization, VITA, in the computer industry, and explained that such agreements would be judged under a rule of reason in which the efficiencies of advance negotiations were recognized (U.S. Department of Justice, Letter to VMEbus International Trade Association, 2006).

**Remedies**

Suppose Firm A unknowingly infringes Firm B’s patent in a way that makes a trivial use of the patent, but it would take Firm A one year to alter the product design. Should B be allowed to threaten to shut A down? Such a threat will allow B to extract A’s profits for the year. It also will give B an incentive to encourage A to go forward initially without revealing B’s patent. Under the principles governing equitable relief, a court may be able to decide whether to allow exclusion or to allow A to pay a “reasonable” royalty (eBay Inc. v. MercExchange, L.L.C., 126 S. Ct. 1837 [2006]). This last approach may be sensible in some circumstances, but it raises a serious danger that courts could become regulators of patent royalties. Limiting the “reasonable” royalty to a design-around period may be one way to mitigate this danger, though one must be wary as to whether Firm A might fail to
make reasonable efforts to gain knowledge about relevant patents, counting on the
court to require only a "reasonable" royalty if its infringement is discovered.

How Should Rules for Exclusionary Conduct Be Modernized?

Alleged exclusionary "bad practices" often promise efficiencies but risk anti-
competitive harms. Deep price cuts benefit consumers, but can also drive rivals out
of business and set the stage for higher prices. Exclusive territories create incentives
for dealers to promote a brand, thereby increasing competition against other
brands, but also eliminate competition amongst dealers of that brand. Tying the
sale of two products together can be a convenience to consumers but can hobble
rivals who produce only one of the products. It is easy to conclude that one should
weigh the costs versus the benefits for each potentially exclusionary practice, but
implementing such a test can be exceedingly difficult. That is why the use of safe
harbors is sensible.

In thinking about such tests, I would require that the test consider the
economywide costs and benefits of finding a violation of antitrust law. Even if one
does correctly find a violation for one particular case, will the effect be to chill the
use of an efficient practice in other circumstances? If so, the finding of liability
could impose large costs and the wiser course may be to not attack the practice. Of
course, not attacking a practice that may be in some cases anticompetitive will
create incentives for additional anticompetitive actions. Ultimately, an assessment
of safe harbor rules versus a more activist policy requires an empirical judgment
based on the strength of market forces to correct errors of inaction and on the
likely costs of errors of intervention (Easterbrook, 1984). It would be too extreme
to conclude that the antitrust laws should entirely ignore exclusionary conduct
because there are well-known examples, especially involving exclusive dealing,
where the possibility of harm is well understood.

Consider the problem of tying and bundling, one of the areas of greatest
current concern in exclusionary conduct because these practices are so widespread.
Indeed, most products can be thought of as bundles of characteristics—shoes with
shoelaces, cars with a motor and a radio. In a tying arrangement, good A can be
purchased only if good B is also purchased. With bundling, separate goods can be
purchased, but combined purchases are offered on more attractive terms. Both
courts and economic thinking remain somewhat confused on this topic.8

The key issue in tying and bundling lies in a distinction between price
discrimination and harm to the competitive structure resulting in higher prices.
Tying can be an effective pricing strategy to extract consumer surplus. For example,
consider a setting where the willingness to pay for one product, like salt machines,
is directly related to the intensity of use of some complementary, competitively

---

8 As a starting point for recent literature on tying and bundling, see Whinston (1990, 2001), Tirole
(2005), Nalebuff (2005), Carlton and Waldman (2002, 2005), and the references cited in these papers.
produced product, like raw salt, where the users of a lot of raw salt value the machine more than users of a small amount of salt, as in the case of *International Salt Co. v. U.S.* (332 U.S. 392 [1947]). A monopolist of salt machines will be able to extract some surplus from consumers of salt machines by requiring consumers of salt machines to buy all their raw salt from the monopolist at above competitive prices (assuming raw salt is produced competitively), while leaving unchanged the price of raw salt to others who have no demand for salt machines. Consumers of salt machines likely are worse off, while producers are better off. In fact, such price discrimination probably adds to efficiency.\(^9\) Price discrimination is rampant in the economy and is typically not an antitrust violation, so opposing price discrimination just because it happens through tying seems odd. Moreover, rules about pricing strike again at the heart of competitive activity. Therefore, I would regard all pricing activity dealing only with a single firm’s extraction of consumer surplus as falling in a safe harbor of protected behavior.

Matters are quite different when the tie-in sale alters the shadow prices faced by the market. Consider a tie-in sale in which, in order to buy A, one must also buy B. Say that the production of B exhibits scale economies, and some firms produce only B, while others produce both A and B. In this case, a tie-in might favor firms that produce A and B, and force firms that produce only B to exit the market. This insight explains how a tie can alter industry structure for B and thereby harm consumers of B who do not consume A (Whinston, 1990). Consider the following example due to Robert Gertner. Imagine that the only hotel on a resort island ties the sale of rooms to meals at its restaurant. There are other restaurants on the island where the native workers eat. If as a result of the tie, demand is so reduced at local restaurants that they go out of business, then the hotel can monopolize the sale of meals to natives. When this insight is extended to a dynamic model, it shows how a firm like IBM or Microsoft can use an initial monopoly to maintain monopoly as goods evolve technologically (Carlton and Waldman, 2002). IBM used its market power in mainframe computers to tie its own peripherals, preventing other rivals from succeeding in peripherals whose existence could have encouraged either substitutes for mainframes or other mainframe entrants. Similarly, Microsoft used its market power in its Windows operating systems to tie its browser, thereby hampering developments of other browsers that themselves could have become a platform for alternatives to the Windows operating system.

The recent decision in *LePage’s Inc. v. Minn. Mining & Mfg. Co.* [3M] (324 F.3d 141 [3d Cir. 2003, en banc]) has created concern about whether courts are handling bundling correctly. LePage’s makes private label transparent tape, as does 3M. 3M decided to bundle Scotch Tape, its premier tape brand, with its own private label brand when selling to retail stores. LePage’s sued and claimed an antitrust violation. The court ruled in favor of LePage’s since LePage’s, without a brand

\(^9\) However, Posner (2001) notes that economists too readily accept that price discrimination will lead to efficiency. The profits from discrimination induce the use by firms of resources to engage in it and by consumers to avoid it. Such a use is a waste.
name comparable to Scotch Tape, could not compete in the same way as 3M. This opinion contains numerous flaws (Rubinfeld, 2005): for example, the court does not appear to have examined the profitability of the discount program, nor whether the discount drove price below cost. However, the central flaw is that LePage’s continued to compete, and therefore, the availability of private label transparent tape to non-3M customers was likely unchanged. That fact alone should have ended the inquiry. Moreover, even if discounts may be anticompetitive in a few settings, general attacks on discounts are likely to wind up harming the competitive process.

Some courts, economists, and government agencies have suggested a “no economic sense” or “profit sacrifice” test for any type of exclusionary conduct, in which one asks whether, setting aside the strategic effect on rivals, the questioned action makes sense (Werden, 2006; Melamed, 2006; see Popofsky, 2006, and Salop, 2006, for criticisms of these tests). Although proponents of such tests are sophisticated and might utilize them reasonably, I am skeptical that one test can work well for all the types of exclusionary conduct. Moreover, the logic of the tests cuts to the core of competitive activity. In general, public policy should encourage firms that want to invest in activities that consumers value in order to gain future sales from their rivals. However, because such actions by definition reduce present profits, a blind application of a “profit sacrifice” test could condemn almost any competitive behavior. When a test could potentially challenge a wide array of core competitive behaviors, it becomes dangerous.

**Mechanism Design for the Antitrust Legal Process**

Legal rules affect incentives for the law to be enforced. In the case of antitrust law, key rules that affect incentives include who can sue, what damages are recoverable, who pays attorneys’ fees, and how settlements should reduce damages when some firms settle but others do not. Let us discuss each in turn. The economic answers do not always comport with current laws.

**Who Can Sue?**

A private party that has been injured by anticompetitive acts can sue and recover treble damages plus attorneys’ fees. If a cartel of steel manufacturers raises the price of steel, then a carmaker would be a direct purchaser and someone who purchases a car containing steel would be an indirect purchaser. It is no defense for the steel producers to say that the direct purchasers were able to pass the overcharges along to others, as the U.S. Supreme Court held in *Hanover Shoe, Inc. v. United Shoe Machinary Corp.* (392 U.S. 481 [1968]). In other words, the cartel of steel manufacturers could not claim that the overcharge of $1 to direct purchasers is not damage, because direct purchasers responded by raising their prices to final car consumers.
Under federal law, only a direct purchaser can sue for damages. In *Illinois Brick Co. v. Illinois* (431 U.S. 720 [1977]), the U.S. Supreme Court banned recovery by indirect purchases because 1) it leads to duplicative recovery, and 2) it is hard to track all indirect purchases. The court held that indirect purchasers can seek injunctions, but not damages. However, 25 states and the District of Columbia have passed "Illinois Brick repealer" statutes allowing indirect purchasers to sue in state court, and several other states permit the same result by judicial interpretations (Cavanagh, 2004). Therefore, every federal antitrust suit involving direct purchasers can be accompanied by duplicative and separate state actions for indirect purchasers.

The logic underlying the federal law is that direct purchasers have the most knowledge about an antitrust harm and their incentive to bring suit would be diluted if their damage award were reduced by what they were able to pass on to indirect purchasers (Landes and Posner, 1979). The logic underlying recovery by indirect purchasers rests on two arguments. First, when indirect consumers are harmed by higher prices, it is unfair that they receive no compensation. Second, direct purchasers may be reluctant to damage a business relationship by suing their suppliers, in which case there could be a suboptimal level of deterrence if only direct purchasers can sue.

The first justification, compensation for those indirectly harmed, is not without merit but ignores the importance of deterrence. A legal regime with the appropriate level of deterrence is valuable, even if a harmed indirect purchaser receives nothing. If suits by indirect purchasers lead to suboptimal deterrence because they deprive direct purchasers of an incentive to sue, then there is a justification to ban such suits.

The second concern, that buyers may not sue their suppliers, seems quite real. For example, in the recent Microsoft case, computer distributors that were direct purchasers brought no private cases—which suggests a role for actions initiated on behalf of indirect purchasers. One approach is to allow indirect purchaser suits under federal law only where the direct purchasers from whom they bought have chosen not to sue in sufficient volume, and to preempt state laws allowing suits by indirect purchasers. Though this procedure has complications, it avoids duplicative recoveries and duplicative state trials by indirect purchasers. Finally, even if one fails to resolve the current conflict between state and federal law by preempting state laws allowing indirect purchasers to sue, consolidation of direct and indirect suits into one trial proceeding would be superior to the current situation where separate state and federal trials occur.

**Damages**

Antitrust laws largely ignore the lessons that economics has to teach about optimal damages. The economic theory of damages shows that the optimal penalty equals the expected net harm imposed on the rest of society by the anticompetitive

---

10 I served as an expert opposed to Microsoft.
act divided by the probability of detection (for example, Landes, 1983; Gavil, Kovacic, and Baker, 2002, pp. 1040–46). If all anticompetitive acts were detected and fined, the optimal penalty from, say, an industry cartel, would equal the lost consumer surplus from the price increase from cartelization. Courts use the overcharge on actual purchases as a measure of damages, but typically ignore the loss in surplus from purchasers who cut back or eliminate their purchases in the face of the cartel price increases.

Under the Sherman Act, antitrust damages are set at treble the overcharge. A multiple of three is appropriate if we detect anticompetitive actions only one-third of the time. However, anticompetitive activities surely vary a great deal in how likely they are to be detected. For example, cartel activity is typically covert and thus hard to detect, so some multiple for damages seems appropriate. It is hard to estimate the current detection probability of cartels, though it has presumably risen in light of recent government leniency programs to encourage cartel members to confess in combination with the recent use of larger fines and longer jail sentences. In contrast, a firm’s decision to employ exclusive territories, or to use bundled discounts to tie its product to others, or to lower its price are all observable in the marketplace. When the acts are observable to all, then single damages (which would include lost social surplus) become optimal. For overt exclusionary practices, the current damage system probably over-deters.

Complications also arise in setting damages when certain conduct such as price fixing has international effects. Courts have generally held that purchases by foreign buyers in foreign countries are not subject to the jurisdiction of the U.S. antitrust laws so that, for example, in the recent case of an international cartel to fix the price of vitamins, foreign buyers of foreign vitamins could not sue in the United States in F. Hoffmann-La Roche Ltd. v. Empagran S.A. (542 U.S. 155 [2004]). If antitrust damages are set based on U.S. damages, but similar damages or other penalties are not available outside the United States, then U.S. antitrust damages will not optimally deter an international cartel. Having U.S. courts award damages to U.S. customers based on international damages is one solution, but it would raise problems of international comity in which U.S. courts would be viewed as interfering with the right of other countries to decide how to regulate their own economies.\footnote{An additional complication arises because increasing damage awards can disrupt the operation of U.S. and foreign leniency programs in which government prosecutors grant a complete or partial reduction in government penalties to firms that reveal a cartel and/or that provide evidence to prosecute a cartel.}

Perhaps the best approach here is to continue the task of convincing other countries to develop and enforce their own strong laws against cartels.

Claim Reduction

Most antitrust lawsuits are settled, rather than litigated to conclusion. However, the threat of litigation influences the settlement terms. Suppose that A and B form a cartel and raise prices by 10 percent on total purchases of $10 million. Thus,
the overcharge of $1 million is multiplied to $3 million in treble damages. Each cartel participant is jointly and severally liable, which means that if B has no assets, A would have to pay the full $3 million to the plaintiffs—and vice versa. If B settles for $1 million and A goes to trial and loses, A is liable for $3 million minus a claim reduction of the $1 million settlement, or $2 million, even if B has a much larger share of the market than A.

In this setting, the plaintiffs are indifferent to how the total damages are split between A and B. Indeed, plaintiffs may want to reward the first to settle by accepting a low settlement, since the first to settle often provides valuable evidence against the cartel. Realizing this, A and B will in effect bid for the right to be first to settle. This process can lead to an equilibrium in which the aggregate amount of settlement exceeds the optimal amount of damages, leading to overdeterrrence (Easterbrook, Landes, and Posner, 1980). One solution would be to make nonsettlement defendants jointly and severally liable for damages based on their collective market share of sales, though this rule would likely increase the use of judicial resources since the incentive to settle could be diminished.

**Antitrust versus Regulation**

Antitrust and regulation represent two alternative approaches to competition policy. Antitrust is designed to let markets work when they can work. Regulation is specific, setting rules for prices and quantities. When markets fail—as in natural monopolies—antitrust is not a substitute for regulation. The U.S. Supreme Court drew a sharp line in *Verizon Commc'ns Inc. v. Law Offices of Curtis V. Trinko, LLP* (540 U.S. 398, 407-08 [2004]): Antitrust courts have no business imposing affirmative duties on a firm to sell at specific prices, with rare exceptions. Where prices must be set by other than the market, then that is a task for regulators.

My position is that regulation should be confined to as few areas as necessary, because history has revealed that regulation, even well-intentioned, can wind up leading to inefficiencies as regulators set policies designed to please various interest groups. Antitrust, when administered by judges not beholden to special interests and when guided by economic reasoning, has shown itself to be a valuable tool to promote efficiency. Indeed, in Carlton and Picker (2007), my coauthor and I argue that antitrust has been developing a comparative advantage to regulation in many settings over the last 30 years, and this has led to a decline in regulation and increasing reliance on the antitrust laws to control competition.

An interesting intersection of antitrust and regulation occurs where there is an “antitrust savings clause,” as in the Telecommunications Act of 1996, which ensures

---

12 The exception is *Aspen Skiing Co. v. Aspen Highlands Skiing Corp.* (472 U.S. 585 [1985]), where the Court ruled it an antitrust violation for one firm to cease cooperating with the other when in the Court’s view the cessation of cooperation was harmful to consumers. For a criticism of *Aspen Skiing*, see Carlton (2001). I served as an expert for Verizon in *Trinko*. 
that even the regulated industry remains subject to the antitrust laws. Thus, the antitrust authorities at the Department of Justice can challenge a merger between two phone companies, even if the regulators at the Federal Communications Commission have approved it. In this way, antitrust and regulation can be viewed as complementary, with regulation confined to as few areas as necessary and with antitrust covering the remainder.

Although regulation and antitrust are administered by separate agencies in the United States, other countries often combine these roles. In Australia, the same agency responsible for enforcing the antitrust laws is also in charge of regulating certain industries. A good research question is whether regulatory capture is less likely when those who enforce the antitrust law are the same as those who regulate industries, or more generally, whether market performance in such industries is better with a single government overseer than with two.

**Conclusion**

Economics has had an enormous positive effect on the evolution of antitrust policy over the last 30 years or so. However, the evolving forces of technology and globalization, together with experience gained over time, suggest that further modernization is in order, though the principles underlying antitrust remain sound.

This paper has suggested a number of lessons. The definition of markets should be de-emphasized, even in the horizontal merger context where it remains somewhat useful, but especially in looking at exclusionary practices and markets where new products and technology are important. Antitrust law can be adjusted to help intellectual property strike the right balance between encouraging innovation and hindering it, with adjustments to assure that anticompetitive patent settlements are discouraged, standard setting is encouraged, and remedies for short-term patent infringement are not draconian. Regulation of exclusionary “bad practices” should be adjusted to establish delimited safe harbors for core competitive behavior like discounting prices of products sold separately or in bundles, introducing products, and entry. The incentives to pursue anticompetitive behavior can be brought in line with the social benefits through several steps: eliminating state laws allowing suits by indirect purchasers, or at least coordinating the lawsuits of direct and indirect purchasers; reducing the multiple on damages for openly observed “bad acts”; avoiding a rush to settle cases too rapidly and for too much; and clarifying the complementary tasks of antitrust and regulation in regulated industries.

Clearly, antitrust offers a rich array of theoretical and empirical issues. For the United States, no comprehensive study yet exists that quantifies the benefits (or costs) of our current antitrust policies compared to other possible policy regimes. (For a debate concerning whether the benefits of antitrust outweigh its costs, see the exchange in this journal between Crandall and Winston (2003) and Baker...
Studies of how individual cases have affected the specific firms involved in the case, though useful, are not a substitute for examining the overall economic effects of the rules in a particular case on all firms. For the rest of the world, the question is whether countries with relatively limited experience with antitrust—that is, most of the world—will learn from past mistakes of U.S. antitrust policy, or repeat them.

This article reflects my views alone and not necessarily those of the U.S. Department of Justice. I thank Thomas Barnett, Norman Familant, Kenneth Heyer, James Hines, Alvin Klevorick, Sheldon Kimmel, William Landes, Ann Norman, Andrei Schleifer, Timothy Tardiff, Michael Waldman, Hill Wellford, and Gregory Werden for helpful comments.

References


