The Antitrust Revolution
To the memory of my father, John Kwoka, and my friend, Bill Burnett.
J. E. K.

To David, the best of all sons.
L.J.W.
Contents

Preface xi
Contributors xv
About the Editors xix
Introduction 1

PART I. Horizontal Structure

The Economic and Legal Context 8

FRANK A. WOLAK
SHAUN D. MCRAE
Page 30

R. PRESTON MCAFEE
DAVID S. SIBLEY
MICHAEL A. WILLIAMS
Page 67

PATRICK DEGRABA
Page 89

RICHARD J. GILBERT
JAMES RATLIFF
Page 115

MARGARET E. GUERIN-CALVERT
Page 140

JONATHAN B. BAKER
Page 157

Page vii
Serdar Dalkir  
Frederick R. Warren-Boulton

PART II. Horizontal Practices

The Economic and Legal Context

Kenneth G. Elzinga  
David E. Mills

John P. Bigelow  
Robert D. Willig

Gary L. Roberts

John M. Connor

Robert H. Porter  
J. Douglas Zona

Part III. Vertical and Related Market Issues

The Economic and Legal Context

BARRY NALEBUFF

Michael L. Katz

BARRY NALEBUFF

F. M. Scherer
Case 17. Revisiting Maximum Resale Price Maintenance: 456
State Oil v. Khan (1997)
GUSTAVO E. BAMBERGER

PART IV. Network Issues

The Economic and Legal Context 474

DENNIS W. CARLTON
HALSIDER

ROBERT S. FINDYCK

DANIEL L. RUBINFELD

JEFFREY K. MACKIE-MASON
JOHN METZLER
The term *revolution*—literally a "turnaround"—has come to be applied to a wide variety of events, products, and ideas. This book is about the truly revolutionary transformation of modern antitrust into an economics-based policy. The idea that economics should be central to antitrust policy is not new, but for much of its history antitrust was driven by wider objectives. Over the past forty years, however, economics has made steady progress in influencing antitrust policy, to the point that many now see antitrust as an extension of economics.

This book chronicles this revolutionary transformation. It consists of recent antitrust cases written by economists who were involved in them. These case studies provide insight into how economists think about antitrust issues as well as how economics now influences the entire process: which cases are brought, how they are evaluated, and how they are presented in courts of law. To those ends, each case provides a detailed description of key issues, arguments, and evidence. Each provides an evaluation of the economic and legal significance of the proceeding. And each explains the effect of the case on the companies, industry, and practices that it covers. In all of these respects the cases reflect the increasingly central role for economics.

The first edition of *The Antitrust Revolution*, twenty years ago, was gratifyingly well received. Many reported to us the enthusiasm of students, instructors, and practitioners of antitrust for the type of economics-oriented case studies that comprises this volume. That reception encouraged us to pursue a subsequent edition, then more, and now this, the fifth. In one sense our motivation has been unchanged throughout: to capture and convey the ever-greater role of industrial organization economics in the antitrust process. But over time this purpose has taken on other dimensions.

For example, it has become apparent that the "antitrust revolution" is an ongoing phenomenon. There is no universally agreed-upon target toward which antitrust analysis and policy are moving. Rather, the underlying industrial organization economics is an evolving body of understanding, so new ways of looking at issues are constantly emerging. Each new edition of *The Antitrust
Revolution reflects these changing perspectives, as it should. In addition, however, these five editions collectively make clear that economics plays a central role in a growing number of cases, and certainly in all major antitrust cases of our time.

This edition is much changed from its predecessors. Of the twenty-one cases, twelve are entirely new. These new cases reflect the increased frequency with which economically important issues are being raised and decided. These newly written cases include several involving merger policy, an area that appears to be undergoing substantial change. These merger cases include the proposed Exelon-PSEG electricity merger, Arch Coal, Oracle-PeopleSoft, and the cruise lines. This edition also includes a number of conduct-related cases, ranging from predatory conduct and so-called reverse payments to exclusive dealing and tying, bundled rebates, and network questions. Two cases also involve intellectual property. While many of these concepts are familiar, their current manifestations have raised new issues or issues that are now being examined with better economic analysis. These are very much a part of our revolution.

As before, while most of the cases in this volume involve antitrust issues raised before the Federal Trade Commission or the Justice Department’s Antitrust Division, they also concern cases brought by private parties, as well as others pursued by the states’ attorneys general, or raised before regulatory bodies in their competition-protection role. The authors represent in some cases the sides that prevailed, and in others the sides whose arguments fell short. And while most of these cases are resolved, at least a couple remain in the courts or are subject to further appeal.

We have kept several cases from the last edition where those cases continue to illustrate important issues and applications of industrial organization economics. As always, we regret not being able to retain more previous cases, but Oxford University Press has posted all previous cases not in this edition—back through the first edition—on its website, so that instructors, students, and practitioners can access them conveniently. These can be found at www.oup.com/us/antitrustrevolution.

Sharp-eyed readers of the past editions will notice that within each of our four sections we have now arrayed the order of the cases such that the most recent cases are presented first (rather than the cases’ appearing in chronological order, as was true in the earlier editions).

For this fifth edition we would like to express out gratitude to the many people who have provided assistance. These include Terry Vaughn and Catherine Rae, our editors at Oxford University Press, for their support for this project. In addition, we want to thank our numerous authors—now thirty-one in number—for their interest in participating in this project, for their willingness to write balanced accounts of cases about which they invariably feel strongly, and for responding to all our suggestions and deadlines.
Most of all, however, we want again to thank our own students, students everywhere, instructors, and practitioners who continue to read *The Antitrust Revolution*. Your enthusiastic reactions and your helpful comments have contributed much to making this fifth edition a reality. We hope that this volume meets your expectations.

J. E. K.
L. J. W.
Contributors

Jonathan B. Baker is a professor at the Washington College of Law, American University. He has served as director of the Bureau of Economics of the Federal Trade Commission, senior economist at the Council of Economic Advisers, and special assistant to the deputy assistant attorney general for economics in the Antitrust Division of the U.S. Department of Justice.

Gustavo E. Bamberger is a senior vice president at Lexecon. He has testified as an expert witness in antitrust and merger cases in the United States, New Zealand, and Canada.

John P. Bigelow is a senior economist at the Princeton Economics Group in Princeton, New Jersey. He has served as an expert witness in antitrust matters in federal court and was formerly a member of the economics department at Yale University.

Dennis W. Carlton is a professor in the Graduate School of Business at the University of Chicago, an NBER research associate, and recently served as deputy assistant attorney general for economic analysis at the U.S. Department of Justice. He has been a member of the economics departments of MIT and the University of Chicago as well as the University of Chicago Law School and served as a commissioner on the Antitrust Modernization Commission.

John M. Connor is professor of industrial economics at Purdue University and an adviser to the American Antitrust Institute. He was formerly head of food manufacturing research at the Economic Research Service of the U.S. Department of Agriculture.

Serdar Dalkir is a Ph.D. economist and a principal at Microeconomic Consulting and Research Associates.

Patrick DeGraba is a staff economist at the Federal Trade Commission. Past positions include principal at Charles River Associates, deputy chief economist of the U.S. Federal Communications Commission, and assistant professor at Cornell University’s Johnson Graduate School of Management.
Kenneth G. Elzinga is Robert C. Taylor Professor of Economics at the University of Virginia. He was one of the first special economic assistants at the Antitrust Division of the U.S. Department of Justice and more recently was Thomas Jefferson Visiting Scholar at Cambridge University.

Richard J. Gilbert is professor of economics at the University of California, Berkeley. He served as deputy assistant attorney general for antitrust in the U.S. Department of Justice, where he led the effort that developed joint Department of Justice and Federal Trade Commission “Antitrust Guidelines for the Licensing of Intellectual Property.”

Margaret E. Guerin-Calvert is president of Competition Policy Associates. She has worked on antitrust matters since 1979 in both public (as assistant chief at the Antitrust Division, Department of Justice, and economist at the Federal Reserve Board) and private sectors, working on merger and nonmerger matters in a wide variety of industries, such as banking, transportation, and health care.

Michael L. Katz is the Harvey Golub Professor in Business Leadership at New York University’s Stern School of Management and also holds the Sarin Chair in Strategy and Leadership at the University of California, Berkeley’s Haas School of Business. He served as chief economist of the Federal Communications Commission from 1994 to 1996 and as deputy assistant attorney general for economic analysis in the U.S. Department of Justice from 2001 to 2003.

Jeffrey K. MacKie-Mason is the Arthur W. Burks Collegiate Professor of Information and Computer Science at the University of Michigan, where he is also professor of economics and public policy. He is the senior partner of ApplEcon LLC and has been a national fellow at the Hoover Institute of Stanford University.

R. Preston McAfee is the J. Stanley Johnson Professor at the California Institute of Technology and vice president of Yahoo! Inc. He has written extensively on auctions, pricing, and antitrust.

Shaun D. McRae is a Ph.D. candidate in economics at Stanford University.

John Metzler is the managing partner at ApplEcon LLC.

David E. Mills is professor of economics at the University of Virginia.

Barry Nalebuff is the Milton Steinbach Professor of Economics and Management at the Yale School of Management.

Robert S. Pindyck is the Bank of Tokyo-Mitsubishi Professor of Economics and Finance in the Sloan School of Management at MIT. His research and writing have covered topics in microeconomics, industrial organization, and finance, and he has served as a consultant or expert witness in a variety of antitrust matters.

Robert H. Porter is William R. Kenan Jr. Professor of Economics at Northwestern University and a research associate of the National Bureau of Economic Research.
James Ratliff is a senior vice president at Competition Policy Associates. He has been a member of the faculty of the economics department at the University of Arizona and has taught at the Haas School of Business, University of California, Berkeley.

Gary L. Roberts is vice president at CRA International. Previously he served as associate director for antitrust in the Bureau of Economics at the Federal Trade Commission.

Daniel L. Rubinfeld is Robert L. Bridges Professor of Law and professor of economics at the University of California, Berkeley. He served as deputy attorney general for antitrust in the U.S. Department of Justice during 1997 and 1998.

F. M. Scherer is professor emeritus at the John F. Kennedy School of Government, Harvard University, and continues to teach courses on industrial economics and technological change. Between 1974 and 1976 he was director of the Federal Trade Commission's Bureau of Economics.

David S. Sibley is the John Michael Stuart Centennial Professor of Economics at the University of Texas at Austin. During 2003 and 2004, he was deputy assistant attorney general for economics in the Antitrust Division of the U.S. Department of Justice.

Hal Sider is a senior vice president at Lexecon in Chicago.

Frederick R. Warren-Boulton is a principal of Microeconomic Consulting and Research Associates. He served as chief economist and deputy assistant attorney general in the U.S. Department of Justice and has been a resident scholar at the American Enterprise Institute, an associate professor of economics at Washington University, and a research associate professor of psychology at American University.

Michael A. Williams is a director at ERS Group and was formerly an economist in the Antitrust Division of the U.S. Department of Justice. He has provided expert testimony before U.S. district courts, the U.S. Court of Federal Claims, and federal regulatory agencies, and he has consulted on antitrust issues for the Antitrust Division and the Federal Trade Commission.

Robert D. Willig is professor of economics and public affairs at Princeton University. He is a senior consultant to Competition Policy Associates and served as deputy assistant attorney general in the Antitrust Division of the U.S. Department of Justice.

Frank A. Wolak is a professor of economics at Stanford University. He specializes in the study of privatization, competition, and regulation in network industries such as electricity, telecommunications, water supply, natural gas, and postal delivery services.

J. Douglas Zona holds the position of director at LECG. He has consulted on antitrust and econometric issues for the Antitrust Division, U.S. Department of Justice, and the Federal Trade Commission.
About the Editors

John E. Kwoka, Jr., is the Neal F. Finnegan Distinguished Professor of Economics at Northeastern University. He is also a member of the board of directors of the Industrial Organization Society, a research fellow of the American Antitrust Institute, and a member of ENCORE, an international network of antitrust scholars centered at the University of Amsterdam. He has previously taught at George Washington University and the University of North Carolina at Chapel Hill, held visiting positions at Harvard University and Northwestern University, and served at the Federal Trade Commission, the Antitrust Division of the Justice Department, and the Federal Communications Commission.

Lawrence J. White is the Arthur E. Imperatore Professor of Economics at the Stern School of Business, New York University, and deputy chair of Stern’s economics department. He is also the general editor of the Review of Industrial Organization and secretary-treasurer of the Western Economic Association International. He has taken leave from NYU to serve in the U.S. government three times: During 1986-1989 he was a board member on the Federal Home Loan Bank Board; during 1982-1983 he was the chief economist of the Antitrust Division of the U.S. Department of Justice; and in 1978-1979 he was a senior staff economist on the President’s Council of Economic Advisers.
The Antitrust Revolution
Introduction

Antitrust policy in the United States now spans three different centuries and many epochs in this nation’s economic history. The first law to be enacted—the Sherman Act of 1890—was a reaction to widespread discontent with business during the Industrial Revolution. The Clayton Act and the Federal Trade Commission Act of 1914 were directed at anticompetitive mergers and conduct that persisted as the nation moved to industrial maturity. Most recently, this country has found itself in the midst of an equally profound Information Revolution. Each of these periods has raised questions concerning the effects of dominant firms, mergers, collusive behavior, vertical integration, predatory pricing, tying, and other matters. And each of these periods has answered these questions with an antitrust policy that has reflected the objectives and the understanding of its time.

But over the past thirty to forty years another revolution has taken place—a revolution in antitrust policy itself. This revolution has involved the ascendance of industrial organization economics in antitrust policy making, with profound effects on the institutions, interpretation, and enforcement of antitrust. The concept of antitrust has had broad political and popular support throughout its history, but there never has been an equivalent consensus about what actual policy should be. One reason is that the language of the original laws left many of the important details unresolved, so that policy came into focus only as the courts provided interpretations of such terms as monopolization, substantial lessening of competition, and conspiracy. Initially, that process produced a formalistic interpretation of the law without much regard for the growing body of knowledge about the economics of companies and industries. Another force was the tension between populism—which gave rise to the antitrust laws—and the equally strong belief that private enterprise ought to be left alone. This tension has run through the entire history of antitrust, periodically...
resulting in policy shifts reflecting then-current sentiment or changing views about certain business practices.

These views started to change in the late 1960s. The Antitrust Division of the Justice Department promulgated Merger Guidelines in 1968, which were developed by a panel of economic and policy experts alongside the staff lawyers of the division and embodied an industrial organization framework for analysis. The courts issued some notable opinions that made clear their unease with aspects of their previous approaches to antitrust issues and their receptivity to economic arguments. The Federal Trade Commission and the Justice Department both began having prominent academics as their chief economists or economic advisers. That practice, in turn, brought other economists to the agency staffs, strengthening their economic expertise and ensuring more sophisticated analysis within the agencies.

Over the past three or four decades this revolution has progressed to the point that industrial organization economics now plays a crucial role in virtually all phases of antitrust policymaking. Economics helps determine what cases the Justice Department and the Federal Trade Commission pursue. Economics frames the central issues for investigation and, based on data analysis and theory, evaluates the likely competitive effects of various practices by companies or of structural changes in industries. And the courts themselves have embraced economic reasoning in their own analyses, a change made possible by the increased economic sophistication of the judiciary. All of this progressed to the point that a prominent antitrust scholar declared that while "antitrust is, first and most obviously, law," it "is also a set of continually evolving theories about the economics of industrial organization" (Bork 1978, p. 10).

A bit more history may be helpful. The first salvo in this economic revolution in antitrust was represented by the so-called structure-conduct-performance school of economics, often associated with the work of Edward Mason, Joe Bain, and others trained at Harvard University. This perspective emphasized the structural roots of competition, a view that found support in path-breaking empirical work relating industry concentration to profits and price-cost margins. By extension it looked askance at many mergers, much conduct of dominant firms, vertical integration, and even conglomerate mergers. It is worth emphasizing that the structure-conduct-performance school enjoyed wide support in the profession at the time, finding expression in the first Merger Guidelines of 1968 and in the first appointments of leading economists to the antitrust agencies. The influence of this school on the agencies, the courts, and antitrust policy was a tremendously significant event.

The second wave of the antitrust revolution was "Chicago School" economics, so named for its place of origin, the University of Chicago, where many of its prominent advocates taught, notably, Aaron Director and George
Stigler. Beginning in the 1970s, that school emphasized the use of basic microeconomic theory for evaluating the effects of industry structure and conduct on economic performance. It argued, for example, that mergers should be analyzed in terms of both their likely price effects and the plausibly cost savings achieved by the merged company. It further claimed that price increases are not so easy to achieve, either because of the inherent difficulty of tacit cooperation or because of ease of entry by new competitors. For these reasons mergers were said to be generally pro-competitive, and not properly evaluated with such indicia as market shares and industry concentration. On other issues, the Chicago School was equally adamant. Price cuts were said almost invariably to reflect lower costs and legitimate competitive behavior rather than predation. Efforts by manufacturers to establish retail prices or to constrain the behavior of independent retailers almost always represent efforts to control certain aspects of the sale in which the manufacturers have a legitimate interest.

The different perspective of the Chicago School extended to its view of the very purposes of antitrust. It argued that antitrust should be guided solely by economic efficiency. Efficiency, this school maintained, is what the plain language of the law implies, and in any event is the only objective that can rationally be pursued. And it has documented numerous instances in which the pursuit of other objectives has actually imposed costs on consumers rather than enhancing the competitiveness of markets.

This was—and is—by no means the consensus view either of the original purposes of antitrust or of what antitrust should now pursue, but the challenge represented by the Chicago School both sharpened the focus of antitrust and helped to discredit some of its more dubious past pursuits. For example, most students of antitrust are at some time led through cases of the 1960s that endorsed the populist objective of protecting small business and that prohibited mergers between companies with small market shares. By the 1980s and 1990s, by contrast, mergers of huge petroleum companies were approved with only minor modification, and mergers in the steel and airline industries won approval despite large market shares. In more recent years, numerous mergers in highly concentrated industries have been permitted.

Similar changes in antitrust policy are apparent with respect to firm conduct. Whereas earlier Supreme Court cases held that virtually any tampering with market price was illegal per se, by the late 1970s the Court admitted the possibility of pro-competitive Justifications even for price fixing by horizontal competitors. In the area of predatory pricing, the traditional view seemed to be that price cuts injuring competitors were evidence of predation, but more recent cases have adopted a permissive view of what is considered acceptable pricing behavior on the part of an incumbent firm. And in contrast to earlier hostility toward most vertical mergers and price agreements, the antitrust agencies now challenge few such arrangements. Also gone from the agenda of the agencies are cases involving price
discrimination, with its generally ambivalent economic effects; potential competition, partly because the judicial standard of proof is so high and partly because of the view that potential entrants are numerous; and conglomerate mergers, previously challenged on potential competition grounds or simply on the basis of their sheer size.

These changes have not resulted merely from the Chicago School critique alone, since many economists of all persuasions had long pressed for antitrust policy that better reflected evolving economic understanding. That said, many believed the Chicago School approach to be too simplistic and dangerously close to repealing much of antitrust. These economists argued that market shares and concentration were informative—if not dispositive—about competitive conditions, and that entry was rarely so quick, cheap, and easy as to obviate concerns about cooperative behavior among existing companies. They raised serious reservations about permissive policies with respect to price cutting and other dominant firm practices, arguing that predatory or disciplining behavior does indeed occur. A significant number of economists were unwilling to go so far as to absolve vertical relationships of all anticompetitive potential. And many rejected the contention that strict economic efficiency was or should be the essential purpose of antitrust.

Over the past twenty years, this counter to the Chicago School approach has been advanced through the infusion of more advanced theory and empirical work into antitrust economics. More powerful theory, better adapted to specific issues, has proven capable of identifying specific conditions under which various practices may have anticompetitive effects, even if they are elsewhere benign. Careful consideration of information imperfections, sunk costs, reputation effects, and strategic behavior has improved our understanding of many matters, including predatory pricing, vertical restraints, tying and bundling, and raising rivals’ costs. In addition, techniques of empirical analysis have become much more sophisticated, with data better suited to the task, models well grounded in theory, and superior econometric tools. Empirical evidence, for example, has given new support to the proposition that concentration affects competition and pricing within industries, a contention eroded by earlier critiques.

There has been, in short, a reconsideration of the entire range of antitrust issues. This more nuanced “post-Chicago economics” argues that many formulations of the preceding twenty years were reliant on overly simplistic theory, with the result that important distinctions were overlooked and excessively sweeping conclusions were drawn. It contends that many practices must be evaluated in light of facts specific to the case rather than being pigeonholed into theoretical boxes. And it is more skeptical of the ability of the market automatically to discipline firms and thereby negate the anticompetitive potential of mergers and various practices.

Post-Chicago economics is not—or at least not yet—a unified alternative paradigm. It has not displaced the Chicago approach in many quarters.
And some have expressed concern that its more fact-based approach will make determinations of antitrust violations more difficult—a possibility that is underscored by some recent court decisions that seemingly require an extraordinary degree of proof of anticompetitive effects before taking action in specific cases. But it has gained acceptance as an intellectually rigorous alternative approach to antitrust. And, of course, it is very much a part of the economic revolution in antitrust. Economics constitutes its foundation just as much as the economics of their times guided the structure-conduct-performance school and the Chicago School. These new views simply represent another step in that revolution. There no doubt will be many more such steps, as economics strives to clarify the effects of structural changes and various business practices on market performance.

While the outcome of that process cannot be foretold, two predictions can safely be made. The first is that the paramount importance of economics in the antitrust process is firmly established. Enforcement policy and court decisions will be grounded in economic analysis to an ever-greater degree. Supporters and critics of policy issues all now debate them in terms of competition and efficiency, clearly conceding the central role that economics plays.

In addition, these advances in economic understanding continually improve the rationality and consistency of antitrust policy. As these advances gain acceptance, they progressively narrow the range within which policy decisions are made. That is, by demonstrating that some propositions are incorrect, lack generality, or suffer from other defects, the advances limit the degree to which future policy can ever revert to those defective propositions.

That does not imply complete agreement about the proper course of antitrust. A considerable range of acceptable policy remains, and there is—and will be—legitimate disagreement over goals and strategies within that range. But to an increasing extent that range is bounded by economics and will shrink as our economic understanding grows. The antitrust revolution is secure.

REFERENCE

PART I
Horizontal Structure
The Economic and Legal Context

Markets may be structured in a wide variety of ways, but purely structural concerns in antitrust arise in the cases of monopolies and oligopolies. Since the monopoly model is at the heart of almost all antitrust analysis, we will begin with a discussion of monopoly and then move on to oligopoly.

MONOPOLY

Economics

The microeconomic theory of monopoly is straightforward: A single seller of a good or service, for which (at a price that would just yield normal profits) there are no good substitutes and for which entry is difficult, will be able to take advantage of its market power. If the seller can sell only at a single price to all buyers (i.e., it cannot practice price discrimination), then its pursuit of maximum profits will lead it to sell a smaller output and maintain a higher price than would an otherwise similar competitive industry.

Figure 1-1 portrays this outcome. As the figure indicates, the maximizing price will not be at "the sky's the limit" levels but instead will be related to the demand curve (via the derived marginal revenue) for the monopolist's output and the monopolist's marginal costs. An immediately important point is that the demand curve—which expresses the empirical reality that at higher prices customers generally buy less—does limit the extent to which the monopolist's price can exceed competitive levels.

This monopoly outcome is socially less efficient than the competitive outcome, because of allocative inefficiency: The monopolist produces too

1 A similar argument applies to monopsony: a single buyer in a market, who can gain by buying less and at a lower price than if competition among buyers prevailed.

2 The familiar formula for profit maximization is \( P = MC/(1 + 1/E_D) \), where \( E_D \) is the (negative) elasticity of demand. Once \( P \), has been determined, the monopolist's output \( Q_M \) can be derived from the demand relationship.
little; equivalently, there are buyers who are willing to buy at prices that are above marginal costs (and who would be able to buy from a competitive industry) but who are not willing to buy at the higher monopoly price (and who buy other, less desirable things instead). The lost consumer surplus of the buyers, portrayed in Figure 1-1, is frequently described as a "deadweight loss triangle."

The higher price (which is the cause of the allocative inefficiency) also yields the higher profits or "overcharge" of the monopolist (which is sometimes described as "monopoly rents"). This overcharge is largely a transfer from buyers to the monopolist and is represented by a rectangle in Figure 1-1.

Since the monopolist should be willing to spend an amount up to the size of the rectangle to defend its monopoly, some of this rectangle may be "burned up" in costly efforts (e.g., political lobbying, raising barriers to entry) to protect its position. Such efforts represent a socially wasteful use of resources and thus add to the deadweight loss of monopoly. See Posner (1975). Also, the absence of competitive pressures may induce less than fully efficient production processes ("X-inefficiency") and thereby add to deadweight loss. See Leibenstein (1966).
Any monopoly seller would like to be able to practice price discrimination, thereby segmenting the firm’s market on the basis of the buyers’ willingness to pay. In order for price discrimination successfully to occur, (1) there must be buyers with different willingnesses to pay; (2) the firm must be able to identify who they are (or have some mechanism that will cause them to reveal themselves); and (3) the firm must be able to prevent arbitrage (i.e., prevent the buyers who receive low prices from reselling to the buyers who would otherwise receive high prices). If the seller could identify each buyer and make an all-or-nothing offer to that buyer at the latter’s maximum willingness to pay, this would constitute “perfect” price discrimination (frequently described as “first-degree” price discrimination). Other forms of price discrimination can involve block pricing (“second-degree” price discrimination), segmenting buyers by geography or by customer type (“third-degree” price discrimination), bundling (Adams and Yellen 1976), and tying (Burstein 1960).

Instances of true monopoly can be found in the U.S. economy, although they collectively account for only a small fraction of U.S. GDP. Examples include local residential telephone service (in some geographic areas), local electricity distribution, local natural gas distribution, postal service for first-class and bulk mail, the single hardware store (or gasoline station, or pharmacy) in an isolated crossroads town, and firms producing unique products that are protected by patents (e.g., those patented pharmaceuticals for which there are no good substitutes). Over time, technological advances tend to erode existing monopolies (e.g., by producing substitutes and by expanding market boundaries through reduced telecommunications and transportation costs), but also to create new ones.

It is a short conceptual leap from the single seller to the dominant firm—a firm of uniquely large size but one that also faces a “fringe” of smaller competitors. Though technically not a monopoly, the dominant firm will still be able to enjoy the fruits of its market power. The extent of its enjoyment will be determined by its costs relative to those of smaller firms, the elasticity of the demand for the product, the elasticity of supply by the fringe, and the ease or difficulty of entry. Historical examples of such market structures include U.S. Steel in steel, Alcoa in aluminum, IBM in mainframe computers, Xerox in photocopying, and Kodak in cameras and film, at least for some time periods. More recent examples include Microsoft in

---

4 This condition is clearly satisfied by a monopoly; it is also satisfied by any firm that faces a negatively sloped demand curve (i.e., that sells a differentiated product), so price discrimination can arise under monopolistic competition or differentiated oligopoly as well.

5 One paradoxical consequence of such perfect price discrimination is that the allocative inefficiency of the monopolist disappears, even while the transfer to the monopolist increases.

6 See Stigler (1965) and Landes and Posner (1981). Though the dominant firm model is usually presented in terms of a commodity product, where the fringe firms’ disadvantages lie in their inferior production technology, the model readily extends to the case of a differentiated product, where the fringe firms’ disadvantages lie in their inferior brand acceptance.
personal computer operating systems, Intel in microprocessors, and United Parcel Service for small package delivery services.

Monopoly can arise in four ways. First, economies of scale may indicate that a single firm is the most efficient structure for serving the entire market. In essence, the technology of production may be such that unit costs decline over the relevant range of production. It is important to note that this "natural" monopoly outcome is dependent on both the nature of the technology and the size of the market. Thus, where markets are small, monopoly may be more likely, whereas larger markets may be able to accommodate multiple efficient producers (if unit costs do not continue to fall at relatively high volumes, or if product differentiation is important to buyers). Also, although engineering relationships often indicate unlimited economies of scale, the difficulties of managing a larger enterprise may yield higher rather than lower unit costs at higher volumes.

Second, incumbent firms may merge to create a monopoly or a dominant firm. Historically, the merger wave of 1887-1904 yielded a large number of such consolidations, including U.S. Steel (steel), Standard Oil (petroleum), American Tobacco (cigarettes), American Can (tin cans), Kodak (cameras and films), DuPont (explosives), and more than sixty other monopolies or dominant firms (Markham 1955; Nelson 1959; Scherer and Ross 1990, ch. 5). As will be discussed below, an important goal of modern antitrust policy is to prevent the creation of market power through mergers (of which a merger-to-monopoly would be the limiting case).

Third, a firm may own a unique and advantageous input into production. For example, market power may arise from the ownership of a unique natural resource (e.g., metallic ores) or the ownership of some patents—e.g., Polaroid's early patents on self-developing film, Xerox's early patents on photocopying, pharmaceutical companies' patents on unique drugs, and Intel's patents on its microprocessors. However, most patents convey little or no market power. All are intended to encourage investment in new ideas and their implementation by creating property rights that prevent quick and easy free-riding on the efforts of innovators. It is this aspect of patents that leads to their description as "intellectual property."

Fourth, government policy can be the source of monopoly. Historically, exclusive government franchises—for rail, air, and trucking service (between some city pairs); local and long-distance telephone service; local cable television service; local banking; and postal service—have yielded monopolies, along with government regulation to deal with them. With the advent of the deregulation movement of the mid-1970s and after, such government-protected monopolies have become more rare, though not wholly extinct.

See Gilbert (1999) and Case 20 by Daniel Rubinfeld in Part IV of this book.
A monopoly or a dominant firm may be able to entrench or enhance its position by raising barriers to entry or raising the costs of its rivals (Salop and Scheffman 1983, 1987). Such efforts will be the subject of discussion in many of the cases in Parts II, III, and IV of this book.

Antitrust

The primary efforts of government to deal with monopoly have been through explicit regulation—for example, through formal regulatory commission or boards—or through government ownership. But from its beginnings in 1890, antitrust law has tried to address monopoly issues. Section 2 of the Sherman Act creates a felony offense for "every person who shall monopolize, or attempt to monopolize, or combine or conspire with any other person or persons, to monopolize . . . ."

The antitrust approach to horizontal structural issues, however, has not been especially potent, at least since 1920. Two important Supreme Court cases in 1911—Standard Oil and American Tobacco—yielded government victories and the structural dissolution of dominant firms in the petroleum and tobacco industries, and similar government victories followed for the next nine years. The Court’s reasoning in those cases, however, established a "rule of reason" that applies to monopolization cases: Courts should consider behavior and intent and efficiencies, as well as just monopoly structure. This approach led to the government’s loss in 1920 in U.S. Steel, after which the government became wary of bringing such cases.

A renewed vigor in antitrust enforcement in the late 1930s led to a suit against Alcoa, yielding a final appellate decision (U.S. v. Aluminum Company of America, 148 F.2d 416 [1945]) that was perhaps the high-water mark in emphasis on structure—but also the turning point. Declaring that Alcoa's 90 percent market share of aluminum clearly represented monopoly, the court appeared to stand ready to infer monopoly from its high share. Yet the court also went on to state other reasons why Alcoa should be convicted, and those reasons were in essence its bad acts. The ambiguity over whether structure or conduct was key was not resolved until the Grinnell case (U.S. v. Grinnell Corp., 384 U.S. 563 [1966]), in which the Supreme Court stated that a violation of Section 2 required two factors: possession of monopoly power and willful acts to acquire or maintain such power.

The effect of this language was to require detailed examination of acts and practices of monopoly firms for their effects, perhaps for intent, and for alternative explanations. All of these have rendered most such proceedings

---

9 U.S. v. Standard Oil Co. of New Jersey et al, 221 U.S. 1 (1911).
"Sec. 2, after all, condemns the effort to ”monopolize,” not the structure of monopoly.
extremely long, complex, and too often unclear. Together with judicial reluctance to tamper with firm structures, this has resulted in few subsequent cases where the government has sought structural relief, and still fewer successes. Government-initiated monopolization cases have usually focused on behavioral remedies, as have privately initiated monopolization cases."

In sum, antitrust (at least since 1920) has not played a large role in dealing with monopoly market structures through horizontal structural relief. Absent a major change in the legal environment, this will likely continue to hold true.

**OLIGOPOLY**

**Economics**

The essence of oligopoly is that the number of sellers is few enough so that each seller is aware of the identity of its rivals and aware that its own actions affect their decisions (and that the others probably have similar perceptions). This condition is sometimes described as "conjectural interdependence."

A wide range of price-quantity outcomes is theoretically possible. At one extreme, a tightly disciplined cartel may be able to maintain prices and quantities that approximate those of a monopoly; at the other extreme, if sellers myopically focus on price competition in a commodity industry, only two sellers are necessary to approximate the competitive outcome."

Accordingly, there is no definitive price-quantity "solution" or outcome for an oligopoly market structure (unlike the specific outcomes that can be predicted for a monopoly structure and for perfect competition).

As will be discussed more thoroughly in Part II, economic theory argues that market structure characteristics (e.g., the number and size-distribution of sellers, conditions of entry, the characteristics of the sellers and of their products, and the characteristics of buyers) are likely to influence the ease or difficulty with which sellers can come to a mutually beneficial understanding with respect to prices or other important dimensions of conduct. In turn, this will imply differences in market outcomes."

Thus, there are

"A survey can be found in Scherer and Ross (1990, ch. 12). The government's success in achieving a 1982 consent decree that broke up AT&T involved vertical structural relief; see Noll and Owen (1994). Similarly, the government's short-lived remedy in its victory over Microsoft involved vertical structural relief; see Case 20 by Daniel Rubinfeld in Part IV.

"See the case discussions in Parts II, III, and IV of this book. Also, private plaintiffs legally cannot obtain structural relief from the courts; see International Telephone & Telegraph Corp. v. General Telephone & Electronics Corp. et al. 518 F.2d 913 (1975).

"However, product differentiation "softens" the competition and results in a less competitive outcome.

"In addition to the discussion in Part II, see the overviews provided by Shapiro (1989) and, especially with respect to horizontal mergers. Jacquemin and Slade (1989).
important links between oligopoly structure and conduct; oligopoly structure matters—which brings us naturally to the consideration of antitrust.

**Antitrust**

The primary vehicle for a structural approach to oligopoly is Section 7 of the Clayton Act, which instructs the Department of Justice (DOJ) and the Federal Trade Commission (FTC) to prevent mergers "where in any line of commerce or in any activity affecting commerce in any section of the country, the effect of such acquisition may be to substantially lessen competition, or to tend to create a monopoly." Though the Act was passed in 1914, Section 7 was largely a dead letter until 1950 because of an unintended loophole.

The Celler-Kefauver Act of 1950 closed that loophole, and Section 7 gained life.

A series of government challenges to mergers in the 1950s and 1960s led to a set of important Supreme Court decisions, beginning with *Brown Shoe* in 1962. In those decisions the Court indicated that it was ready to prohibit both horizontal mergers between competitors and vertical mergers between customers and suppliers, even in markets where the merging parties' shares were relatively small and entry was easy. The Court expressed concerns about competition but also opined that Congress had intended to halt mergers so as to preserve market structures with large numbers of firms, even at the sacrifice of some efficiency that might be achieved by a merger. However, the Court backed off from this tough, semipopulist position in two merger decisions in 1974. The Supreme Court has not rendered a decision on a government challenge to a merger since then.

Flushed by the favorable Supreme Court decisions of the 1960s, the DOJ’s Antitrust Division developed a set of Merger Guidelines in 1968. The Guidelines indicated the circumstances (described in terms of industry four-firm concentration ratios and the sales shares of the merging firms) in which the DOJ would be likely to challenge mergers, so that the private antitrust bar could provide better guidance to its clients. While those Guidelines reflected the economic and policy understanding of the time, it soon became apparent that they were too restrictive and too rigid. As a result they fell into disuse during the 1970s, awaiting an effort to revise them in accordance with advances in economics and different views about appropriate policy.

17 Parts II, III, and IV of this book will address behavioral approaches to addressing oligopoly issues.

18 The 1914 act forbade mergers that were effected through one company’s purchase of another company’s equity shares. Merger candidates quickly realized that they could easily evade this restriction by simply having one company buy all of the underlying assets of the other company.


The 1968 Guidelines were largely scrapped, and a new set was issued in 1982. Economists played a large role in the development of the new Guidelines and in subsequent revisions in 1984, 1992 (when the FTC joined as an author), and 1997. These Guidelines have proved influential in shaping antitrust lawyers', economists', and eventually judges' approaches to mergers (Werden 2003). They certainly shaped many of the economic arguments that were developed in the cases discussed in this part. Accordingly, we next turn to a more detailed discussion of the Guidelines.

**THE HORIZONTAL MERGER GUIDELINES**

The Horizontal Merger Guidelines start from the fundamental premise that the antimerger provisions of the Clayton Act are intended to prevent the exercise or enhancement of market power that might arise as a consequence of a merger. They thus reject the populist position that the pure sizes of the merging entities should be a consideration in the evaluation of a merger.

Using the analytical base of the microeconomics of monopoly and oligopoly discussed above and in Part II, the Guidelines address six crucial issues:

- the delineation of the market for merger analysis, so as to determine whether the merger partners compete with each other and the sizes of their (and other relevant sellers') market shares;
- the level of seller concentration in a relevant market that should raise antitrust concern about a merger;
- the potential adverse effects of mergers, either through coordinated behavior among sellers or through the possibility that the merging firms might unilaterally, postmerger, be able to affect prices and output;
- the extent and role of entry into the market;
- other characteristics of market structure that might make the postmerger exercise of market power easier or more difficult; and
- the extent to which merger-related cost savings and efficiencies that are promised should be allowed as a defense of a merger that appears to

"Also, the 1992 revision modified the title to Horizontal Merger Guidelines.

increase the likelihood of the exercise of market power, and the types of
efficiency evidence that should be considered.

Each will be discussed in turn.

Market Definition
The Guidelines define a relevant market for antitrust merger analysis as a
product (or group of products) sold by a group of sellers who, if they acted
in concert (i.e., as a "hypothetical monopolist"), could bring about "a small
but significant and nontransitory increase in price" (SSNIP). This is equiv­
alent to defining a relevant market as one in which market power can be
exercised (or one in which existing market power can be enhanced). The
Guidelines indicate that a 5 percent price increase sustained for one year is
the likely SSNIP value that the enforcement agencies will use. The smallest
group of sellers that satisfies the SSNIP test is usually selected as the rele­
vant market. These principles apply to the determination of both product
markets and geographic markets. Under this definition, markets might be as
small as a neighborhood or as large as the entire global economy; the deter­
mining factor is simply whether buyers would switch in sufficient numbers
to undermine the price increase.

The logic of this approach follows from the goal of preventing merg­
ers that create or enhance market power as measured by such a price
increase. The SSNIP test identifies the smallest group of sellers who could
exercise such market power. With one exception, the market definition par­
adigm focuses on sellers (since it is sellers who exercise market power).
That exception arises when a group of sellers could practice price discrim­
ination and raise prices significantly for an identifiable group of customers
(defined by a geographic area or a business function). In such a case, that
group of customers may also be considered to be a relevant market.

A stylized graphical example, using geographic differentiation, can
further illustrate these points concerning market definition. In Figure 1-2,
imagine a set of, say, shoe stores (A, B, C, . . .) stretched along a highway,
with one mile between each store. Customers are located continuously
along the highway in between the stores. Initially, let us assume that any
customer will buy only from the lowest-price store location that is immedi­
ately to her right or left; that is, she will not "skip over" a store to get to
the next one. (Let us call this a strong preference for adjacency.) Each store
effectively competes with the stores on each side of it and gets a roughly
equal share of the potential customers on each side of it.

\[The \ example \ uses \ geographic \ space \ and \ one \ dimension; \ but \ the \ insights \ apply \ equally \ well \ to \ product \ space \ and \ to \ multiple \ dimensions.\]

\[For \ the \ sake \ of \ simplicity, \ we \ abstract \ from \ transportation \ costs.\]
If stores D and E merge, then the customers located between them are "trapped"; if the merged D-E firm can identify which of its customers are from the "trapped" region and can price discriminate, it will likely raise its prices to them. If the price increase satisfies the SSNIP criterion, the merger will be anticompetitive, and the trapped customers between any two adjacent stores will thus constitute a relevant market. Even if the merged D-E firm could not identify the trapped customers (or devise a means whereby they revealed themselves), it might still find that it could raise its overall prices to considerable heights and earn sufficiently huge profits from its trapped customers (if they had quite inelastic demands for shoes) to offset its lost profits on the lost sales to the nontrapped (former) customers (its former share of customers between C and D and between E and F). In this case the merger would be anticompetitive, and any two adjacent firms would constitute a relevant market.

Now let us suppose that the buyers' preferences for adjacency are less strong, and customers are willing to skip over one store in their pursuit of cheaper shoes. Effectively, each store competes with the three stores to its right and the three stores to its left. The merger between stores D and E would no longer pose a competitive problem, because none of their customers is trapped.

Further, in this case the minimum number of adjacent stores that could possibly constitute a relevant market would be four. (Two stores clearly do not, and even three do not, since there are still no trapped customers.) If a hypothetical monopolist were to control stores C, D, E, and F, customers between D and E would be trapped. If the monopolist could practice price discrimination, it could profitably raise the price to them, and the customers in the middle of the four stores constitute a relevant market. If price discrimination is not possible, then it seems unlikely that the gains from charging high prices to the trapped D-E customers would offset the lost profits on the lost sales to the (former) nontrapped customers (all of the customers between C and D and between E and F, and its share of the customers between B and C and between F and G). So, the relevant market would have to include a larger number of adjacent stores (such that the hypothetical monopolist's gains from enough trapped customers more than offset its loss of profits from the lost nontrapped customers). Further analysis (of the type described below) would be necessary to determine whether a merger

\[\text{This might also be considered an instance of "unilateral effects," which is discussed below.}\]
between any two sellers within this market (of, say, any adjacent seven stores) would create an anticompetitive problem.”

As this example illustrates, the strength of buyers’ preferences for individual companies’ products vis-à-vis other individual companies’ products is key to the delineation of the market.

**Seller Concentration**

With the market boundaries determined, the analysis turns to the postmerger level of seller concentration and the merger-induced change in that concentration that would trigger enforcement attention. The conventional basis for this approach is the belief that cooperative conduct is more likely at higher seller concentration levels (see, for example, Stigler 1964), although an alternative theory of anticompetitive harm—the “unilateral effects” approach (which will be discussed below)—is also considered.

The *Guidelines* use the Herfindahl-Hirschman Index (HHI) for this measurement. The HHI for a market is computed by summing the squared market shares (expressed as percentages) of all of the sellers in the market. Thus, an atomistic market would have an HHI very close to zero; a pure monopoly would have an HHI of 10,000 ($100^2 = 10,000$); and a duopoly consisting of two firms with, for example, 70 percent and 30 percent market shares, respectively, would have an HHI of 5800 ($70^2 + 30^2 = 5800$).

The *Guidelines* specify two nominal decision points: For an “unconcentrated” market with a postmerger HHI below 1000, the merger will rarely, if ever, be challenged. For a “highly concentrated” market with a postmerger HHI above 1800, if the merger itself causes an increase in the HHI of 100 or more, there is a presumption that the merger is anticompetitive; whereas, if the increase in HHI is between 50 and 100, there is heightened scrutiny of the merger. In either event, other factors (e.g., ease of entry, strong buyer power) could overcome this presumption. For a “moderately concentrated” market with a postmerger HHI between 1000 and 1800 and a merger-based increase of 100, the presumption of competitive concern is weaker.

“This example also points out an interesting paradox. Suppose that the seven sellers B, C, D, E, F, G, and H were found to constitute a market when customers were willing to skip over one store. A conspiracy among the seven sellers to raise prices would be unlikely to stick, since the flanking firms (B and H) would lose customers and receive no gains and thus would likely leave the conspiracy, thereby exposing C and F, etc. But if E owned B and H, then the flanks would be protected. Thus, a three-way merger of E with B and H might be more conducive to a coordinated price increase by the seven than would the merger of E with D and F.

“A quick method of determining the change in the HHI that is due to the merger of two firms is to multiply their premerger shares and then double the result.

“There are two ways of translating the HHI threshold points into more familiar terms. An HHI of 1000 would be yielded by a market with ten equal-size firms (each with a 10 percent market share); an HHI of 1800 would be yielded by a market of between five and six equal-size firms. Alternatively (since most markets do not have equal-size firms), the two decision points translate
In practice, the enforcement agencies have been considerably more lenient than the nominal HHI thresholds would indicate. Rarely have mergers in postmerger markets with an HHI of less than 2000 been challenged, and mergers in markets with substantially higher postmerger HHIs have also escaped challenge. In essence, the merging parties in such instances have been able to convince the enforcement agencies that other characteristics of the market and/or the merger make the postmerger exercise of market power unlikely.

**Adverse Effects**

The *Guidelines* present two theories—coordinated effects and unilateral effects—concerning the adverse effects of mergers. The first and more traditional approach holds that a heightened probability of coordinated behavior would arise as a consequence of an industry structure with fewer sellers and with the merged firm having a larger market share. The *Guidelines* recognize that characteristics other than seller concentration can affect the likelihood of coordinated behavior, including entry (which is discussed in the following section) and other features of the market that could facilitate monitoring and policing of any seller understandings (which are discussed in other market characteristics).

Alongside the traditional concern that postmerger oligopolistic sellers might coordinate their behavior, the *Guidelines* (since 1992) mention a second mechanism that could result in competitive harm: unilateral effects. Unilateral effects arise in markets where, even in the absence of cooperation with other sellers, the merged firm could find a unilateral price increase profitable. The most obvious circumstance in which this might occur is when the two merging sellers are each other’s major competitors in a differentiated product market, so that the elimination of competition between the two as a result of the merger significantly relaxes the prior pricing constraint that each felt. Other products are simply too imperfect as substitutes to prevent price increases, and for the same reason other sellers’ cooperation in the price increase is not necessary.

Empirically (on the basis of simple correlations) to four-firm concentration ratios of approximately 50 percent and 70 percent, respectively (Kwoka 1985).

"This leniency (as compared with the Guidelines’ formal HHI benchmarks) has been an "open secret" of merger enforcement for over two decades. Recent data released by the FTC and DOJ have confirmed what was suggested far earlier by Lodish (1986). See, for example, FTC and DOJ (2003), FTC (2004), Kwoka (2004), Coate (2006), and Coate and Ulrick (2005)."

"Enforcement procedures now almost invariably set forth one of these two possible concerns as part of agency challenges to a merger.

"The Guidelines highlight a postmerger market share of 35 percent for the merging firm, along with strong customer preferences for the two premerger firms’ products, as worthy of special concern."
Ordover and Willig (1993) provide a stylized example of unilateral effects, which we adapt for this discussion. Suppose that all frozen beets are produced by a single company, B. Another single company, C, produces all frozen carrots, and all frozen spinach is produced by company S. Each company has set its own prices so as to maximize its own profits. An important constraint on each company's pricing is whatever elasticities of substitution among the different vegetables exist among consumers.

Now suppose that beet producer B merges with carrot company C. Two things now occur. First, the merged company BC would now find that a higher price for beets would be profitable, because some of the lost customers switch to carrots, and so the merged company BC gains (internalizes) some profits from those customers, which the stand-alone company B would not have gained; similarly, a higher price for carrots is profitable for BC where it was not profitable for the stand-alone company C. Second, the merged company BC can do better than this outcome because it is now able to set the price of both products simultaneously so as to maximize joint profits.

Note that the competitive harm from this merger does not result from collusion or cooperation; rather, it results from the ability of the merged company to internalize more of the benefits of the price increase. Clearly, this result depends heavily upon the premerger pattern of elasticities and cross-elasticities among all products in the differentiated product setting. While these are not always easy to estimate or specify, antitrust enforcement has made some progress in this area over the past few decades. Aided by detailed scanner data (data recorded at point of purchase in supermarkets, drug stores, and the like), the FTC and DOJ (as well as defendants) often estimate models of differentiated product competition in the context of actual mergers.

It is also worth noting that the Guidelines market definition paradigm is essential (and was developed) for the coordinated effects theory of adverse consequences of a merger, since seller concentration—and thus the delineation of a market within which seller market shares can be calculated—is central to that approach. However, the market definition paradigm fits less well—indeed, it may well be redundant—in the context of the unilateral effects theory, since that approach rests crucially on the empirical estimates of the premerger elasticities and cross-elasticities among a group of differentiated products (some of which are being merged). In essence,

"Ordover and Willig’s example has more subtleties in that they allow for multiple producers of beets and of carrots, and hypothesize a merger between leading producers of each. While perhaps more realistic, the issues are much the same as those set forth here.

"See Werden and Froeb (1994).

For discussions of these techniques, see, for example, Werden and Froeb (1994) and Shapiro (1995). This kind of approach was used in the analysis of the Oracle-PeopleSoft merger, which is discussed in Case 2 by Preston McAfee, David Sibley, and Michael Williams in this part.
if these empirical estimates show that the merged firm would likely find unilateral price increases profitable, then this result automatically indicates that one or more relevant markets (that encompass the products for which postmerger price increases are likely to be profitable) are present."

Entry

Since easy entry by new firms could thwart sellers' efforts to exercise market power even in highly concentrated markets, the Guidelines recognize entry as an important component of merger analysis. They recognize that, for entry to obviate concerns about the potential for postmerger exercise of market power, it must be "timely, likely, and sufficient in magnitude, character, and scope." Timeliness requires entry to occur within a period of two years. The criterion of likelihood is satisfied if the entrant would be profitable in the post-entry market. Sufficiency in magnitude, character, and scope requires that the entrant be capable of restoring the degree of competition that is lost as a result of the merger—that is, that the entrant be as capable and vigorous an entity as the one eliminated by the merger."

The Guidelines acknowledge that high levels of "sunk costs" can be a significant barrier to entry; "sunk costs" are the acquisition costs of tangible and intangible assets that are "uniquely incurred to serve the relevant... market" and that cannot be completely recouped by redeploying them elsewhere. (Examples include specialized production equipment, marketing costs, training costs, research and development, advertising, etc.) The Guidelines specifically ask whether, despite the presence of sunk costs, sufficient entry would be likely to occur within two years in response to a merger-induced price increase."

Other Market Characteristics

As mentioned earlier, the traditional theory of postmerger seller coordination recognizes that other market characteristics can influence the market outcome. Sellers always have an incentive to "cheat" on any implicit (or explicit) understanding among themselves that tempers their competition, especially if they believe that such cheating (e.g., price cutting) can go undetected for a considerable period. Accordingly, the ability of sellers to

"See, for example, White (2006).

"It is interesting to note that the current Guidelines do not set out an explicit method for analyzing "potential entrants": firms that are poised to enter the market quickly. Such firms can exercise significant constraint on existing sellers, making a market more competitive than is apparent from a consideration of existing sellers alone, but also raising significant competitive issues if the potential entrant is itself a party to the merger. For further discussion, see Kwoka (2001).

"Firms that could enter easily (i.e., without the expenditure of significant sunk costs) within one year are considered to be in the market, as part of the market delineation process."
detect and "punish" (through, e.g., severe price cutting) deviations from any understanding is important for the success of any sustained period of noncompetitive behavior.

The Guidelines discuss the major market characteristics that oligopoly theory recognizes as important determinants of sellers' abilities to detect and punish deviations and thus to coordinate their behavior:

- the availability to all sellers of key information about market conditions and individual transactions;
- typical pricing or marketing practices by firms in the market;
- the level of concentration on the buyers' side of the market;
- the degree of complexity in the quality and service dimensions of the product or products at issue; and
- the antitrust history of the sellers in the relevant market.
Cost Savings and Efficiencies

In principle, the cost efficiencies achieved by a merger could yield social savings that would more than compensate for the social loss created by the exercise of market power. Figure 1-3, drawn from Williamson (1968), illustrates the trade-off. Suppose that a merger converts a competitive industry into a monopoly but also achieves cost efficiencies. The social gain is represented by the rectangle of reduced costs; the social (deadweight) loss is the triangle. If the area of the rectangle exceeds the area of the triangle, the merger yields a net social gain. The overcharge rectangle (which is a transfer from buyers to sellers) may still be an obstacle to a merger if the goal of antitrust is considered to be solely to help consumers or if, as seems to be the case, consumers matter more than do producers.

If, however, the cost reduction is great enough, the postmerger price could be lower than the premerger price, even taking into account the post-merger exercise of market power. Or, as is sometimes argued, the post-merger efficiencies may change the dynamic within a sluggish oligopolistic industry and allow the merged firm to challenge the industry leader aggressively. More often, however, some modest efficiencies may result from a merger, leaving the enforcement agencies and/or the courts with the task of making a judgment about the extent of the possible price increase that might be risked in order to achieve cost savings.

An important practical problem is that efficiencies are easy to promise before a merger, but often difficult to deliver after the fact, especially in the context of the difficulties of the postmerger firm’s efforts to integrate personnel, equipment, systems, and cultures from the two premerger firms. The Guidelines recognize the trade-off and the dilemma and try to strike a compromise, stating that the agencies “will not challenge a merger if cognizable [i.e., merger-specific and verifiable] efficiencies are of a character and a magnitude such that the merger is not likely to be anti-competitive in any market.”

MERGER ENFORCEMENT PROCEDURES

Under the provisions of the Hart-Scott-Rodino Act of 1976 (as amended in 2001), the parties to all prospective mergers that exceed specified thresholds must notify the FTC and DOJ of their intentions to merge and
provide basic information about the companies involved. Within a few days the FTC and the DOJ decide which agency will be responsible for reviewing the merger. The basis for this allocation is usually the presence of expertise about the industry within the agency (although political "horse-trading" between the agencies when prominent cases arise is not unknown). Most mergers receive a quick screening and are found to be innocuous. In instances where there is potential for anticompetitive effects, a group of lawyers and economists within the relevant agency is assigned to undertake further analysis.

The agency has thirty days (fifteen days if the merger involves a hostile takeover) from the initial notification during which the merger cannot be consummated (unless the agency's quick screening reveals an absence of problems and the agency so notifies the parties involved). At the end of that period the parties can consummate their merger unless the agency makes a "second request" for more information. In this event, after the parties deliver the requested information, the agency has an additional thirty days (ten days for a hostile takeover) to reach a decision. In complex and controversial mergers, however, the elapsed time from the initial notification to the agency's decision can be considerably longer than the sixty days, partly because of the parties' delays in delivering requested information and partly because of the agency's requests for delays.

If the agency concludes that a merger does pose a potential problem, the parties and the agency will try to determine if there is an acceptable remedy, or "fix," that would alleviate the agency's competitive concerns and still allow the merger partners to gain the efficiencies or other advantages that they seek from the merger. Typically, solutions can be found whereby the merger partners sell off facilities (for some product lines and/or in some geographic areas) to smaller rivals or to entrants, so as to reduce the relevant HHIs to acceptable levels. For example, in mergers between large banks with overlapping branch networks in multiple metropolitan areas, a standard remedy is to require the merging banks to sell sufficient branches to smaller rivals so as to decrease the HHI levels in each metropolitan area to acceptable levels."

"The act was the response to complaints by the enforcement agencies that they sometimes found out about mergers late or even only after the event and that legally "unscrambling the eggs" of a completed merger created an unnecessary extra burden on merger enforcement.

"The parties' lawyers often request meetings with agency officials to present their case for the absence of competitive harm, to which they typically bring company executives and economics consultants/experts.

"This was true, for example, of the BP-ARCO merger discussed by Bulow and Shapiro (2004).

""An interesting analysis of these remedies can be found in FTC (1999).
If an acceptable remedy cannot be found, the agency will indicate its intention to challenge the merger in court. Often, this announcement alone will cause parties that are unwilling to endure the additional delays, costs, and uncertainties of a court challenge to abandon the merger.  If they choose to contest the agency’s action, the agency will typically seek and quickly obtain a temporary restraining order from a federal district court judge.

The agency then asks for a preliminary injunction (PI). Usually, within a few weeks the judge conducts a small-scale trial, lasting a week or two, that is nominally about the fairness of granting a PI but is really a mini-trial on the merits of the two sides’ arguments about the potential anticompetitiveness of the proposed merger.

The judge’s decision on the motion for the PI is often determinative: If the agency wins, the parties are unwilling to appeal and simply cancel the merger;  if the parties win, the agency drops the case. But appeals to a federal circuit court of appeals by either side are possible.  Or the losing party can (but only rarely does) request a full-scale trial on the merits of the case, which can take many months or even years of pretrial maneuvering, extensive document requests and depositions, and a lengthy trial itself.

It should be noted that not all mergers are reviewed by the DOJ or FTC, or even subject to the Merger Guidelines standards. In regulated industries, primary antitrust authority often rests with the regulatory agency,  or authority is shared with the DOJ or FTC.  The regulatory agencies usually evaluate mergers under a broader “public interest” standard, of which antitrust concerns constitute only one part.

WHITHER MERGER ANTITRUST POLICY?

When the "modern" Merger Guidelines were introduced in the early 1980s, they were controversial, inspiring extensive criticism and even alternative

"This happened, for example, in the MCI WorldCom-Sprint proposed merger discussed by Pelcovitz (2004).

"This happened in the Staples-Office Depot proposed merger discussed in Case 7 by Serdar Dalkir and Frederick Warren-Boulton in this part.

"This happened in the Heinz-Beech-Nut proposed merger discussed in Case 6 by Jonathan Baker in this part.

"If the DOJ is the prosecuting agency, the trial takes place in federal district court, and the losing party can then appeal to a circuit court of appeals and then to the Supreme Court. If the FTC is involved, the case is adjudicated by an administrative law judge (ALJ), who then reaches a decision and writes an opinion. The losing party can then appeal to the full Commission for a final agency decision. If the merging parties are unhappy with the Commission’s decision, they can appeal to a circuit court of appeals.

"This was true of the UP-SP merger discussed by Kwoka and White (2004).

"This was true of the proposed EchoStar-DirecTV merger discussed in Case 4 by Richard Gilbert and James Balfiff in this part and the proposed Exelon-Public Service merger discussed in Case 1 by Frank Wolak and Shaun McRae in this part.

Page 25
frameworks. More than a quarter-century later, they have clearly stood the test of time as an organizing framework for the antitrust analysis of mergers. As was discussed above, some later additions and amendments have helped them adjust to subsequent theoretical and empirical advances.

Within the framework, however, criticism still arises, especially with respect to enforcement. On the one hand, some critics believe that too many anticompetitive mergers (e.g., by merger partners with large market shares in industries with high seller HHIs) are being approved by enforcement agency officials who are too sanguine with respect to the prospects of entry and of promised efficiencies and who are too skeptical of the possibilities for coordinated behavior among sellers in high concentration industries. On the other hand, there are those who believe that the agencies are not giving sufficient weight to the prospects of efficiencies that can be created by mergers. The debate on these issues is likely to persist for as long as merger policy remains a feature of antitrust; a resolution that is satisfactory to both sides seems highly unlikely. This debate also highlights, however, the continuing importance and relevance of merger policy in the U.S. economy.

REFERENCES


See, for example, American Antitrust Institute (2004).

This, for example, was prominent in the Antitrust Modernization Commission’s (2007, ch. I.B.) discussion of merger policy.


Kwoka, John E., Jr. "Some Thoughts on Concentration, Market Shares, and Merger Enforcement Policy." Paper presented at the FTC/DOJ Workshop on Merger Enf...
presentations/040217kwoka.pdf.


CASE I

Frank A. Wolak* and Shaun D. McRae

INTRODUCTION

On December 20, 2004, the Exelon Corporation (Exelon) and the Public Service Enterprise Group Incorporated (PSEG) announced a merger agreement to create Exelon Electric and Gas; the parties claimed that it would be the largest utility in the United States. The merged entity would serve close to seven million electricity customers and two million natural gas customers in Illinois, New Jersey, and Pennsylvania and would own slightly more than 51,000 megawatts in generation capacity, with 40 percent of it powered by nuclear energy.

Exelon’s expertise in operating nuclear generation facilities and PSEG’s ownership of 3500 megawatts (MW) in nuclear capacity were cited by the parties as an important factor driving the merger. The merger agreement was accompanied by a separate operating services contract that

*Frank A. Wolak assisted the U.S. Department of Justice in its analysis of the proposed merger. The views expressed herein are not purported to represent those of the U.S. Department of Justice.

See http://phx.corporate-ir.net/phoenix.zhtml?c=124298&p=irol-newsArticle&id=652855 for the merger announcement that contains this statement.

See http://phx.corporate-ir.net/phoenix.zhtml?c=124298&p=irol-newsArticle&id=656265 for the merger announcement that contains this statement.
specified that Exelon would manage plant operations at PSEG’s nuclear generation units regardless of the ultimate outcome of the merger.

Exelon owns PECO, Pennsylvania’s largest utility, which serves approximately 1.6 million electricity customers in the city of Philadelphia and surrounding counties and serves 480,000 natural gas customers outside of Philadelphia. PSEG owns Public Service Electric and Gas Company (PSE&G), which serves 2.1 million electricity customers and 1.7 million natural gas customers in New Jersey. The service territories of PECO and PSE&G are adjacent to one another, and a substantial fraction of the generation capacity owned by these firms is located within or near these two service areas. Consequently, the combination of PECO and PSE&G could reduce competition in the market for wholesale electricity to serve electricity consumers in the PECO and PSE&G service areas and in substantial parts of the remainder of the PJM Interconnection, the wholesale market that covers some or all of the states of Pennsylvania, New Jersey, Maryland, Delaware, North Carolina, Michigan, Ohio, Virginia, West Virginia, Indiana, and Illinois.

Until the electricity industry restructuring of the 1990s and early twenty-first century, most mergers in the electricity supply industry did not raise significant antitrust concerns. The industry was composed of vertically integrated geographic monopolies, each of which was responsible for the production, transmission, distribution, and sale of all electricity in its service territory. Each firm had a legal obligation to serve all customers in its service territory at prices set by a state public utility commission. These prices were set to allow the firm an opportunity to recover all prudently incurred costs of serving its customers, including a return on capital invested by the firm’s shareholders. Consequently, many of the usual mechanisms for a merger to create or enhance market power and thus raise prices to final consumers were not available to electricity suppliers because of this state-level regulation of retail prices and service quality.

Electricity industry restructuring replaced cost-based regulation with market mechanisms as the primary means for setting wholesale electricity prices. Mergers between generation unit owners that sell into wholesale

---


4 Exelon also owns ComEd, which provides electricity to 5.2 million customers in northern Illinois, including the city of Chicago.

5 Vertically integrated electric utilities operating in multiple states created significant challenges for the individual state regulatory commissions. Each state could ensure only that all of the prices that it set allowed the firm an opportunity to recover the costs of serving customers within its boundaries. The larger were the number of states in which a vertically integrated electric utility operated, the more challenging was the effort to ensure that the prices set by all of the state regulatory commissions did not allow the firm to earn total revenues in excess of all prudently incurred production costs. For this reason, mergers involving multistate utilities in the former vertically integrated monopoly regime could cause antitrust harm.
markets can raise significant antitrust concerns because a firm that owns a larger share of the total generation capacity available to serve demand has a greater incentive to withhold output to raise the market-clearing price. The California electricity market during the period of June 2000 to June 2001—documented in Borenstein, Bushnell, and Wolak (2002) and Wolak (2003c)—illustrates the potential impact of the exercise of unilateral market power on short-term wholesale electricity market prices.

Combining two electricity suppliers that own a substantial amount of generation capacity in close proximity to one another—which would have been the final outcome of the proposed merger of Exelon and PSEG—could significantly reduce competition in a bid-based wholesale electricity market. The California electricity market experience suggests that small changes in competitive conditions can lead to substantial wealth transfers from consumers to producers as well as significant deadweight losses. The parties recognized the need to remedy any potential competitive harm associated with the merger. In their initial merger application to the Federal Energy Regulatory Commission (FERC), the parties proposed to sell a total of 2900 MW of fossil fuel generation facilities and 2600 MW of nuclear facilities in a virtual divestiture to address the potential economic harm associated with the merger. Around the time that the merger was announced, the editor of a leading industry publication argued that the primary point of debate between the merging parties and reviewing agencies was likely to be over how best to remedy the effect of the merger on wholesale competition in the PJM Interconnection wholesale electricity market (Radford 2005).

A bid-based short-term market that sets day-ahead and real-time wholesale prices for each hour of the day is a defining feature of formal wholesale electricity markets. These markets yield a rich source of data on the willingness to supply of each producer that can be used to simulate alternative postmerger and postdivestiture market outcomes. The history of state-level regulation of the industry provides information on the technical characteristics of the "fleet" of generation units owned by each firm necessary to construct an accurate estimate of its marginal cost function. Bid-based wholesale markets have market rules that are filed with the FERC—the U.S. wholesale market regulator—that specify precisely how the actions of market participants translate into revenues received and costs incurred. This clarity and detail in the market rules make it unnecessary for the economic analyst to make untestable assumptions about how the market

7 See http://phx.corporate-ir.net/phoenix.zhtml?c=124298&p=irol-newsArticle&ID=670685 for the press release proposing these divestiture packages. The concept of a “virtual divestiture” is explained in the text later in this chapter.
operates in simulating postmerger and postdivestiture market outcomes. These three features of wholesale electricity markets allow a rich quantitative analysis of the impact of mergers and proposed remedies with few economic modeling assumptions besides expected profit-maximizing behavior by firms.

The remainder of the chapter presents a graphical discussion of this quantitative approach to merger analysis in wholesale electricity markets and then applies several insights from it to the proposed PSEG and Exelon merger. The next section describes the PJM Interconnection, the various agencies charged with analyzing the competitive impacts of the merger, and the rationale for the merger offered by the parties. The following section describes important aspects of electricity production and bid-based wholesale electricity markets that are necessary to understand the quantitative approach to merger analysis. A graphical approach is used to convey the basic economic theory and intuition for the quantitative analysis. This discussion is followed by a description of the formal merger review process, including the initial proposal of the merging parties and the outcome of the review process at each agency, including the ultimate outcome of the proposed merger and possible reasons for it. The chapter concludes with a discussion of the lessons that future merging parties in a bid-based wholesale electricity market might learn from the outcome of this proposed merger.

INDUSTRY AND INSTITUTIONAL BACKGROUND FOR MERGER ANALYSIS

This section first describes important features of the PJM Interconnection wholesale market and the characteristics of the two merging parties. This is followed by a description of the four agencies reviewing the merger and the position of the merging parties.

Industry Background for Merger Analysis

Figure 1-1 contains a map of the PJM Interconnection, which is the largest wholesale electricity market in the United States in terms of the total megawatts of generation capacity within its boundaries (approximately 165,000 MW) and the level of peak demand (approximately 145,000 MW). The vast majority of generation units owned by PSEG are contained in the region labeled “PSE&G” in Figure 1-1. The majority of generation units owned by Exelon are in the regions labeled “PECO” and “ComEd.” The shared border between the PSE&G (in New Jersey) and PECO (in Pennsylvania) territories suggests that without an appropriate package of
generation unit divestitures, the merged entity could have a greater incentive and ability to increase wholesale electricity prices, particularly in parts of the PJM Interconnection that contain the PSE&G and PECO service territories.

Table 1-1 lists the characteristics of the two merging parties taken from the PSEG website at the time of the merger. As of the end of 2003, Exelon owned 34,467 MW of generation capacity, and PSEG owned 17,117 MW.

FIGURE 1-1 Map of PJM Interconnection

![Map of PJM Interconnection](image)

**TABLE 1-1**
Characteristics of the Merging Parties

<table>
<thead>
<tr>
<th></th>
<th>Exelon 2003</th>
<th>PSEG 2003</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electric Customers</td>
<td>5,100,000</td>
<td>2,000,000</td>
<td>7,100,000</td>
</tr>
<tr>
<td>Gas Customers</td>
<td>460,000</td>
<td>1,600,000</td>
<td>2,060,000</td>
</tr>
<tr>
<td>U.S. Generation Assets (MW)*</td>
<td>34,467</td>
<td>17,117</td>
<td>51,584</td>
</tr>
<tr>
<td>Nuclear Generation (MW)</td>
<td>16,943</td>
<td>3510</td>
<td>20,453</td>
</tr>
<tr>
<td>Assets, Revenues, and Income (billions of dollars)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Assets</td>
<td>$41.9</td>
<td>$28.1</td>
<td>$70.0</td>
</tr>
<tr>
<td>Total Revenues</td>
<td>$15.8</td>
<td>$11.1</td>
<td>$26.9</td>
</tr>
<tr>
<td>Net Income**</td>
<td>$1.7</td>
<td>$0.9</td>
<td>$2.6</td>
</tr>
</tbody>
</table>

*Projected 2004 year end.

**Income from continuing operations.


Page 34
for a combined total of 51,584 MW. Exelon, which specializes in the operation of nuclear generation facilities, owned 16,943 MW of nuclear generation. PSEG owned 3510 MW of nuclear generation, so that (absent divestitures) the combined entity would own 20,453 MW of nuclear generation assets.

The parties estimated that the merged entity would have had total assets of close to $80 billion and annual revenues of more than $27 billion in 2004. As Radford (2005) notes, this would make the new firm more than 25 percent larger in terms of total assets than the second-largest electricity utility in the United States. As Table 1-1 demonstrates, both PSEG and Exelon also serve a significant number of retail natural gas customers. However, the combined entity would have almost three and a half times more electricity customers than natural gas customers.

The natural gas companies owned by PSEG and Exelon would remain vertically integrated utilities subject to state-level retail price regulation following the merger. It appears that the major reviewing agencies did not believe that the merger would increase the difficulty of regulating the retail price of natural gas charged by the utilities owned by the merged entity. The U.S. Department of Justice (DOJ) complaint opposing the merger (as it was originally proposed) stated that "Exelon's merger with PSEG would eliminate competition between them and give the merged firm the ability and incentive to raise wholesale electricity prices . . . " but did not mention harm to wholesale or retail natural gas markets. The FERC investigation of the merger focused almost exclusively on assessing the impact of the combination on wholesale electricity prices (see FERC 2005).

Reviewing Agencies and Standards for Review

Four agencies assessed the competitive impacts of the proposed merger: the federal wholesale electricity regulator (FERC), the state public utility commissions in the states of New Jersey and Pennsylvania, and the DOJ. Both the federal and state regulatory commissions are legally

8 Exelon was formed in October 2000 by the merger of Unicom of Chicago (which owned ComEd) and PECO. A major motivation for that merger was to consolidate the nuclear operations of the two merging parties.


10 This complaint is available at http://www.usdoj.gov/atr/cases/f216700/216785.htm.

"The Illinois Commerce Commission, the Illinois public utility regulatory commission, informed the merging parties that it did not have jurisdiction over the merger and thus that its approval was not required.
required to apply a "public interest" standard, which is different from the standard applied in the Horizontal Merger Guidelines (HMG), in deciding whether to approve the merger. The HMG focus on assessing whether the merger will "create or enhance market power" or "facilitate its exercise."

The public interest standard of the FERC considers three factors to assess a merger: (1) the effect on competition, (2) the effect on rates, and (3) the effect on regulation. The effect on competition is essentially the HMG competitive effects assessment. The effect on rates asks whether the merger will affect the rates for any wholesale power or transmission customer. This assessment is closely related to the effect-on-competition assessment because if the merger lessens competition, then wholesale prices and rates to some customers are likely to rise. The effect on regulation asks whether the effectiveness of federal or state regulation will be adversely affected by the merger.

The Pennsylvania Public Utility Commission and the New Jersey Board of Public Utilities (NJBPU) have slightly different public interest standards. The Pennsylvania standard requires the merger to provide substantial benefits to Pennsylvania consumers and not to result in anticompetitive or discriminatory conduct or the unlawful exercise of market power. The NJBPU public interest standard requires an assessment of the impact of the merger on competition, rates, employees of the affected public utility, and the provision of safe and adequate utility service, in addition to tangible benefits to New Jersey consumers.

From their public statements, the primary competition concern about the merger by state regulators appears to have been the effect of the combination of generation assets owned by PSEG and Exelon on wholesale market prices and the eventual effect on retail electricity prices in their respective states. The Pennsylvania Public Utilities Commission cited the fact that Exelon agreed to freeze PECO’s retail electricity rates through 2010 as a major reason for approving the merger. The NJBPU stated that "the acquisition of PSEG by Exelon would explicitly reduce the number of significant competitors in New Jersey wholesale markets by one . . ." and " . . . absent mitigation or other measures, the currently substantial market shares of each company in the relevant markets raises not merely the potential but rather certainty of significantly higher market concentration and the potential future exercise of market power" (NJBPU 2005). Therefore, based on these public statements and those cited above, the primary competition concern of both state and federal reviewing agencies appeared to be competition in the wholesale electricity market.

Position of the Merging Parties

The parties claimed that the merger would provide significant value to customers and shareholders of both companies. Most of these benefits were the result of the increased scale of operations of the combined entity. This was claimed to lead to improved service reliability and greater earnings predictability. Greater service reliability was also claimed as a consumer benefit. The combination of several regulated utility businesses and large, low-variable-cost, and low-emissions generation businesses—primarily the nuclear generation units owned by the two companies—in the PJM Interconnection wholesale electricity market was claimed to provide consistent profitability and stable cash flow growth. All of these benefits accrue to shareholders and may be the result only of increased opportunities for the merged entity to exercise unilateral market power in the wholesale electricity market, rather than the result of any cost savings.

The one potential source of cost savings was the opportunity for Exelon to take over the ownership of PSEG’s nuclear “fleet” and implement its management and operation practices at these generation units. Exelon operates the largest nuclear generation fleet in the United States and has implemented a successful nuclear performance program: the Exelon Nuclear Management Model. Data provided by the parties in their merger announcement presentation demonstrated that significant capacity utilization increases and nonfuel production cost reductions were possible at PSEG nuclear facilities from adopting Exelon management practices. In this same presentation, the merging parties argued that every 1 percent increase in the capacity utilization of PSEG’s nuclear fleet would result in an additional $12 million in pretax income.

Although these potential cost savings and revenue increases from improved performance of PSEG nuclear facilities were economically significant, working against the claim advanced by the PSEG and Exelon that these synergies could be realized only through a merger was the fact that a separate nuclear operating services contract was signed at the same time as the merger agreement. Under this agreement Exelon supplies senior personnel to manage daily plant operations and implemented the Exelon Nuclear Management Model at PSEG’s Hope Creek and Salem nuclear generation facilities. Further evidence against the need for PSEG and Exelon to merge in order for PSEG to realize improvements in operating efficiency is the fact that Exelon staff ran the PSEG nuclear facilities during


“The two figures on p. 22 of the presentation on December 20, 2004 ([http://media.corporate-ir.net/media_files/Iro/d/124298/pdfs/EXC_PSEG_AnalystPres_122004.pdf](http://media.corporate-ir.net/media_files/Iro/d/124298/pdfs/EXC_PSEG_AnalystPres_122004.pdf)) show noticeably higher capacity factors and lower nonfuel production costs for Exelon versus PSEG nuclear generation units.
2005 and 2006 under the operating agreement and realized significant operating efficiency improvements relative to those achieved the year before the agreement was signed."

Without divestitures the merged entity would control roughly one-third of the generation capacity in the PJM Interconnection. Because of the close geographic proximity of the PECO and PSEG service territories, the merged entity would control a substantially larger fraction of the generation capacity in the portion of the PJM Interconnection—PJM East—that contains these two service territories. There is a large transmission interface that divides the remainder of PJM from PJM East: the Eastern Interface. A substantial number of the low-variable-cost generation units is located west of the Eastern Interface, and these units typically export electricity to PJM East. Consequently, when there is congestion on the Eastern Interface, the merged entity would face competition from a substantially smaller number of generation unit owners.

Based on the publicly stated preconditions for approval of the merger by each of the reviewing agencies, there appears to be disagreement among them over how to ensure that the merger did not harm competition in wholesale electricity in the PJM Interconnection. For example, the FERC required less fossil fuel capacity to be divested relative to the amount required by the DOJ. However, the FERC required virtual divestiture of nuclear generation capacity, whereas the DOJ did not. The NJBPU never stated its final position on the necessary amount of divestiture, but in August 2006 the New Jersey public advocate proposed an additional divestiture of more than 2500 MW of fossil fuel plants beyond the DOJ settlement." Consequently, the specific points of disagreement on how to remedy the competitive impacts of the merger appeared to be (1) the total MW of generation capacity to be divested, (2) the variable costs and location of the MW of capacity to be divested, and (3) whether a virtual or physical assets sale was sufficient to address these competition concerns. The next section introduces the key features of the electricity supply industry that are necessary to analyze the competitive impacts of a merger in a bid-based wholesale electricity market.

**TOOLS FOR ANTITRUST ANALYSIS IN WHOLESALE ELECTRICITY MARKETS**

There are four segments of the electricity supply industry: (1) generation, (2) transmission, (3) distribution, and (4) retailing. These segments differ in

---


the extent to which the technology of production favors supply by a single geographic monopoly.

Generation is the conversion of heat or kinetic energy such as fossil fuel, nuclear fuel, running water, wind, or solar energy into electrical energy. Transmission is the high-voltage transportation of electricity from generation facilities to the local distribution network. Distribution is the low-voltage delivery of electricity to final consumers. Retailing is the sale of electricity to final consumers. The electricity retailer purchases wholesale electricity and pays for the cost to use the transmission and local distribution networks that deliver the electricity to final consumers.

In a wholesale market, all generation unit owners within the geographic confines of the wholesale market compete to provide wholesale electricity to retailers, who then sell it to final consumers using the local distribution network. All generation unit owners and retailers in this geographic area have access to the transmission network at regulated prices. All retailers pay for local distribution network services at a regulated price set by the state public utility commission.

Wholesale electricity markets function like wholesale markets for other products with a number of very important exceptions that are due to the unique characteristics of electricity. First, supply must equal demand at every instant in time and at every location in the transmission network. Second, all electricity produced must be delivered through a transmission network that has finite capacity to transfer electricity between any two locations. Third, the production of electricity is subject to severe capacity constraints in the sense that a generation unit can produce only a finite amount of energy within an hour. Fourth, electricity is very costly to store, so that virtually all of what is consumed must be produced during that same time period. Finally, the real-time demand for electricity is close to perfectly price-inelastic because the retail price charged to virtually all final consumers does not vary with the hourly wholesale price.

A real-time wholesale market operator is required to ensure that all of these technical requirements for the reliable supply of electricity to final consumers are met. Although the details of wholesale electricity market operation are extremely complex, the basic features can be described in simple economic terms. Each day generation unit owners submit to the market operator their willingness to supply energy from their generation

---

17 Wolak (1999) discusses the technology and organization of production in the electricity supply industry and a comparison of the early experience with wholesale electricity markets.


19 A unit with a nameplate capacity of 200 MW can typically produce slightly more than 200 MWh (on the order of 10 percent) within an hour by operating beyond the unit’s minimum variable cost level of output. Operating in this range also significantly increases the risk that the unit will fail.
units during the following day, as a function of the market price for each hour of the following day.

For all wholesale electricity markets in the United States, these willingness-to-supply functions are step functions, with the height of each step equal to the offer price at which the owner of the generation unit is willing to supply an amount of output equal to the length of the step. For example, in the PJM Interconnection, each supplier is allowed to submit up to ten price levels and associated quantity increments for each generation unit each day. The ability to submit different price offers for different levels of output from the generation unit accounts for the fact that the variable cost of producing electricity can vary with the level of output in the current period and in previous time periods. Because each generation facility is typically composed of multiple generation units, and firms usually own many generation facilities, these firm-level willingness-to-supply curves are composed of hundreds of price levels and quantity increments. Exelon and PSEG each own more than one hundred generation units, which implies that each company sets more than one thousand price levels and quantity increments in the willingness-to-supply curves that it submits to the PJM Interconnection wholesale market operator each hour of the day. To simplify the subsequent graphical analysis, for the remainder of this chapter these willingness-to-supply functions and all marginal cost functions are assumed to be smooth.

In the absence of congestion in the transmission network, the market operator can take the willingness-to-supply curve from each generation unit owner and compute an aggregate supply curve for each hour of the following day. A single market-clearing price is then determined by the price at which the aggregative willingness-to-supply curve intersects the aggregate demand for that hour. All generation unit owners with offer prices less than the market-clearing price are obligated to supply the total megawatt hours (MWh) of energy offered at or below this price.

Figure 1-2 plots the firm-level willingness-to-supply curves for a wholesale electricity market with three suppliers. The vertical axis denotes the market price and the horizontal axis the quantity that each firm is willing to supply at that price. Let \( S_1(p), S_2(p), \) and \( S_3(p) \) equal the willingness-to-supply curves of firms 1, 2, and 3, respectively. At a price of $60/MWh, firm 1 is willing to supply 200 MWh, firm 2 is willing to supply 100 MWh, and firm 3 is willing to supply 300 MWh. The aggregate willingness-to-supply curve, \( S(p) \), is the sum of the amounts that the three firms are willing to supply at each possible price. The vertical line denotes the market demand, \( Q_d \), which is equal to 600 MWh. The price at the intersection of the

---

"Wolak (2007) presents evidence from generation unit-level bidding and operating behavior in the Australian electricity market that unit-level marginal costs vary with the level of output in the current and neighboring time periods of the day."
market demand with the aggregate willingness-to-supply curve is the market-clearing price, which is $60/MWh.

According to the market rules, the revenue of each firm is the market-clearing price times the amount that it produces. In this case, firm 1 receives $12,000 in revenues, the market-clearing price of $60/MWh times 200 MWh, the amount it was willing to supply at that price. Firm 1's profits are the difference between this total revenue and its total cost of producing 200 MWh. The revenues and profits of firms 2 and 3 are determined in an analogous manner.

To analyze a proposed merger in a bid-based wholesale electricity market it is necessary to understand how each firm constructs its willingness-to-supply curve that it submits to the wholesale market operator and how that curve enters the aggregate willingness-to-supply curve that determines the market-clearing price. The essential intuition for this process can be conveyed using the standard model of a profit-maximizing monopoly.²

The first step is to construct a measure of firm-level unilateral market power using the market demand and the willingness-to-supply functions of all other firms besides the one under consideration. Let firm 1 be the

²Wolak (2000, 2003a, and 2007) presents a general model of expected profit-maximizing bidding behavior in wholesale electricity markets with step function willingness-to-supply curves that is the foundation for the graphical analysis presented here.
supplier whose expected profit-maximizing willingness-to-supply curve is being constructed and assume for the moment that firm 1 can observe the market demand and the willingness-to-supply curves of firms 2 and 3. This is clearly not the case in reality because all firms must submit their willingness-to-supply curves at the same time, and the value of the market demand is typically unknown when they do. However, this assumption simplifies the presentation and will be relaxed once a number of important concepts have been introduced.

Figure 1-3 depicts the construction of what is called the "residual demand curve" facing firm 1. At each price, firm 1's residual demand is the amount of market demand left for firm 1, given the willingness-to-supply curves of all its competitors. In Figure 1-3, firm 1's residual demand is computed by taking the difference between the market demand and the amounts that firms 2 and 3 are willing to supply at that price. At a price of $60/MWh, firm 2 is willing to supply 100, and firm 3 is willing to supply 300, so the residual demand facing firm 1 is 200, the difference between the market demand of 600 and the sum of the willingness-to-supply quantities of firms 2 and 3. The curve $D_1$ is the residual demand curve for firm 1 that results from applying this procedure for all prices between $0/MWh and $90/MWh.

The theory of profit-maximizing behavior by a monopoly yields the price that maximizes firm 1's profits, given this residual demand curve, which depends on the market demand and willingness-to-supply curves of firm 1's competitors. Figure 1-4 reproduces firm 1's residual demand curve.
from Figure 1-3, $D_r$ and adds the marginal revenue curve, $MR_r$, associated with this residual demand curve, along with firm 1’s marginal cost curve, $MC$. A profit-maximizing monopolist produces at the level of output where marginal revenue equals marginal cost, $MR = MC$, which implies that firm 1 produces 200 MWh. If firm 1 produces 200 MWh, its residual demand curve yields a market-clearing price of $60/MWh. Firm 1’s variable profit for this residual demand curve and marginal cost function combination equals the shaded area above its marginal cost curve and below the market-clearing price of $60/MWh. Firm 1 cannot obtain higher profits at any other price or level of output, given the willingness-to-supply curves of its competitors and the market demand, than it does at this price and quantity pair. For this reason, the price and quantity pair ($60/MWh, 200 MWh) is called the “best-reply” price and quantity pair for firm 1 for the residual demand curve realization $D_r$.

As noted above, the construction of firm 1’s expected profit-maximizing willingness-to-supply function is complicated by the fact that it does not know the actual residual demand curve realization that it will face when it submits this function to the market operator. However, firms are typically able to observe the level of market demand and the willingness-to-supply curves of their competitors after the market closes for the day. Although the aggregate demand for electricity varies considerably across hours of the day, week, or year, it can be forecast very accurately on a day-ahead basis. The physical operating characteristics and availability of all generation units in the wholesale market are usually known to all market participants.
at the time that they submit their willingness-to-supply curves. These facts and the availability of previous market demand and willingness-to-supply curves provide valuable information about the set of possible residual demand curves that each firm might face and the probability that it will face each of these residual demand curves.

A firm’s residual demand curve can be used to construct a summary measure of its ability to raise the market price by its unilateral actions. Figure 1-5 graphs two possible residual demand curves for firm 1, $D_1$, and one that is much steeper, $D_{1'}$. Both curves pass through the price and quantity pair ($60$/MWh, 200 MWh). If firm 1 faced $D_1$, by selling 50 MWh less, reducing its output from 200 MWh to 150 MWh, the market-clearing price would increase to $67$/MWh. If firm 1 instead faced $D_{1'}$, by selling 50 MWh less it would raise the market-clearing price substantially, from $60$/MWh to $90$/MWh. Consequently, a supplier that faces a steeper residual demand curve has a much greater ability to raise the market price by withholding output than does a supplier that faces a flatter residual demand curve.

The steepness of a residual demand curve can be measured in a way that does not depend on the units used to measure prices and quantities. The price elasticity of demand is defined as the percentage change in the residual demand at price $P$ that results from a 1 percent increase in this price:

$$s(P) = \frac{\text{Percentage Change in Residual Demand}}{\text{Percentage Change in Price}}$$

FIGURE 1-5 Form of Residual Demand Curve and Price Increase from Withholding Output
The inverse of this demand elasticity measures the percentage change in the market-clearing price that results from a 1 percent reduction in the firm's output at price \( P \). The absolute value of this inverse elasticity can be thought of as a measure of the ability of the firm to raise the market-clearing price by reducing its willingness to supply electricity."

All of the characteristics of wholesale electricity markets described above tend to make the elasticity of the residual demand curves faced by large suppliers extremely small in absolute value, which implies extremely large inverse elasticities and very large market-clearing price increases from that supplier's withholding a small percentage of its output. Typically, the greater is the share of total generation capacity owned by a supplier, the smaller is the absolute value of the elasticity of the residual demand curve it faces and the greater is its incentive to raise prices through its unilateral actions. Thus, a supplier that owns a large fraction of the total available capacity is likely to be able to raise prices by more than can a smaller supplier, if both suppliers withhold, say, 10 percent of their generation capacity.

There are actions that electricity retailers and state public utility commissions can take to limit the incentive of large suppliers to exploit their unilateral ability to raise market-clearing prices in a bid-based, short-term wholesale market. The most important is the amount of fixed-price forward contract obligations between the electricity supplier and electricity retailers. The next section explains how these fixed-price forward contracts affect the incentive of suppliers to raise prices in the short-term market. For this reason, we believe that they should play a crucial role in analyzing the competitive effects of mergers in bid-based wholesale markets.

**Impact of Fixed-Price Forward Contracts on Supplier Behavior**

To ensure a reliable supply of wholesale electricity at a reasonable price, electricity retailers sign fixed-price forward contracts that guarantee the price at which they can purchase a fixed quantity of electricity. Let \( P_c \) equal the price at which the supplier agrees to sell energy to an electricity retailer and \( Q_c \) equal the quantity of energy sold at that price. This contract is negotiated in advance of the date that the generation unit owner will supply the energy, so that the values of \( P_c \) and \( Q_c \) are predetermined from the perspective of the supplier's behavior in the short-term wholesale market.

It is straightforward to demonstrate that for the same residual demand curve realization, the larger are a supplier's fixed-price forward

---

"Wolak (2003b) computes hourly residual demand elasticities for the five largest suppliers in the California wholesale electricity market to understand changes in their bidding behavior between the summers of 1998 and 1999 and the summer of 2000 when wholesale electricity prices in California rose dramatically."
contract obligations, the lower will be the price that it finds profit-maximizing because the firm earns only the short-term market price on the difference between its actual production and its forward contract quantity. Therefore, the revenue increase from raising the short-term price is smaller when the firm’s forward contract quantity is larger relative to its actual production.

Figure 1-6 reproduces the residual demand curve and marginal cost curves for firm 1 from Figure 1-4. Suppose that firm 1 has a fixed-price forward contract obligation of 100 MWh. Firm 1’s level of actual output still determines the market-clearing price set by its residual demand curve. For example, if firm 1 produces 200 MWh, the market-clearing price would be $60/MWh, as in Figure 1-4. However, this forward contract obligation alters firm 1’s revenues because it receives the $60/MWh market-clearing price for only 100 MWh. The remaining 100 MWh (of the 200 MWh produced) is sold at $P_c$, the price in the fixed-price forward contract obligation.

To determine the change in total revenues to firm 1 from withholding 1 MWh of output with 100 MWh of fixed-price forward contract obligations, Figure 1-6 plots firm 1’s residual demand curve less its fixed-price forward contract obligations, which are effectively sold at $P_c$, the price in the fixed-price forward contract obligation. The marginal revenue curve for $D_1 - Q_c$, the net-of-forward-contracts residual demand curve facing firm 1, is constructed following the standard approach to constructing marginal revenue curves.

FIGURE 1-6 Best-Reply Price and Quantity with Contracts
To convert $\text{MR}_D$, the marginal revenue curve associated with the net-of-forward-contracts residual demand curve, to the marginal curve associated with firm $l$'s output, shift $\text{MR}_D$, to the right by the amount of firm $l$'s fixed-price forward contract forward obligations. This curve is plotted as $\text{MR}_Q$, and comparing it with $\text{MR}$, in Figure 1-4 reveals that it is uniformly higher at every level of output, which implies a point of intersection of $\text{MR}_Q$ with $\text{MC}_j$ at a higher level of output than the 200 MWh shown in Figure 1-5.

The profit-maximizing level of output for firm 1 for the residual demand curve $D$, and fixed-price forward contract obligation $Q = 100$ shown in Figure 1-6 is 224 MWh, which implies a lower market-clearing price of $56/MWh. This figure demonstrates the general result that for the same residual demand curve realization, the larger is a supplier's fixed-price forward contract obligation, the larger is its best-reply output level and the lower is its best-reply price. Extending this logic to the computation of expected profit-maximizing willingness-to-supply curves implies that for the same distribution of residual demand curves, a larger quantity of fixed-price forward contract obligations leads to a greater willingness to supply output by firm 1 at each possible market price."

The elasticity of a supplier's residual demand curve net of its fixed-price forward contract obligations measures its incentive to raise prices in the short-term market. Let $e(P)$ denote this magnitude, which is defined as the percentage change in the difference between the firm's residual demand at price $P$ and its forward contract position brought about by a 1 percent increase in price. If the firm has a positive amount of fixed-price forward contract obligations, then a given change in the firm's residual demand as a result of a 1 percent increase in the market price implies a much larger percentage change in the firm's net-of-forward-contract-obligations residual demand.

For example, suppose that a firm is currently selling 100 MWh but has 95 MWh of forward contract obligations. If a 1 percent increase in the market price reduces the amount that the firm sells by 0.5 MWh, then the elasticity of the firm's residual demand is $-0.5 = (0.5 \text{ percent quantity reduction}) \div (1 \text{ percent price increase})$. The elasticity of the firm's residual demand net of its forward contract obligations is $-10 = (10 \text{ percent net of forward contract quantity output reduction}) \div (1 \text{ percent price increase})$. Thus, the presence of fixed-price forward contract obligations implies a dramatically diminished incentive to withhold output to raise short-term wholesale prices, despite the fact that the firm has a significant ability to raise short-term wholesale prices through its unilateral actions.

"Wolak (2000) uses bid, market outcome, and forward contract quantity data for large suppliers in the Australian wholesale electricity market to demonstrate the sensitivity of a supplier's incentive to influence the short-term market price to the value of its fixed-price forward contract obligations."
In general, $e'(P)$ and $e(P)$ are related by the equation:

$$e'(P) = e(P)(\text{Actual Output/}[\text{Actual Output} - Q_c]).$$

The elasticity of the firm's residual demand curve net of its fixed-price forward contract obligations is equal to the elasticity of its residual demand curve times the ratio of its total output to its output net of forward contracts. This equation implies that any non-zero value of $e(P)$, which quantifies a firm's ability to raise market-clearing prices by its unilateral actions, can be translated into a very small incentive to raise market-clearing prices (a large value of $s_c(P)$ in absolute value) through a large enough value of forward contract obligations relative to actual output.

If all suppliers have significant fixed-price forward contract obligations, then expected profit-maximizing bidding behavior by firms 2 and 3 implies that these firms will submit willingness-to-supply curves with higher levels of output at each possible price than those given in Figure 1-2. The logic used to construct the residual demand curve facing firm 1 described in Figure 1-3 implies that it will now face a much flatter residual demand curve. Figure 1-5 implies that firm 1 will now have a reduced ability to withhold output to raise the market-clearing price because a 1 MWh reduction in output will increase the market-clearing price by less. This logic demonstrates that a high level of fixed-price forward contract obligations for all suppliers in a wholesale market reduces both the unilateral incentive and the unilateral ability of a supplier to exercise unilateral market power.

Because the magnitude of fixed-price forward contract obligations limits the incentive of suppliers to exercise unilateral market power in the short-term market, one might expect a firm with a significant ability to raise short-term prices to avoid signing forward contracts unless it receives prices that yield the same level of profits that it expects to earn from selling in the short-term market. However, if the buyer of a fixed-price forward contract obligation negotiates this agreement far enough in advance of the date of delivery, then a substantially larger set of firms can compete to supply this product, and the buyer can expect to obtain a more competitive forward contract price. Specifically, if the time between signing the contractual agreement and first delivery of energy from the contract is long enough to allow a new generation facility to enter and provide this energy, the buyer of the fixed-price forward contract will face a very elastic supply of forward contracts at the long-run average cost of a new entrant.

\footnote{As noted in Wolak (2003c), this logic explains why the prices for the forward contracts negotiated by the state of California during the winter of 2001 for delivery starting in the summer of 2001 were so high. California had to pay for market power that these suppliers expected to be able to exercise in the short-term market during the following two years.}

\footnote{The supply of new fixed-price forward contracts is unlikely to be perfectly elastic because certain generation locations or technologies have a finite capacity for expansion. However, the differences...}
Therefore, at this lengthy time horizon to delivery, existing suppliers are unable to raise forward contract prices above the long-run average cost of a new entrant. Moreover, because of efficiency gains in electricity production over the past twenty-five years, there are many existing generation units with variable costs that are far above the variable cost of the production technology employed by a new entrant.

Therefore, if buyers purchase the vast majority of their energy requirements in fixed-price forward contracts far enough in advance of delivery, they can avoid prices that are the result of significant unilateral market power in the forward contract or short-term market. If all buyers follow this advance purchase forward-contracting strategy, and an existing generation unit owner fails to sign a fixed-price forward contract, then it faces a significant risk of selling its output in a short-term market at very low prices because of the high levels of fixed-price forward contract obligations of the remaining suppliers and the additional new generation units constructed to meet the forward contract obligations sold by new entrants. To avoid this unprofitable outcome in a world where all buyers are purchasing the vast majority of their energy needs far in advance of the delivery date, existing generation unit owners should be willing to sign fixed-price forward contracts at prices that are slightly below the long-run average cost of a new entrant."

The above logic illustrates a very important point in the analysis of any horizontal combination in a wholesale electricity market: If the premerger forward contract obligations of the two merging parties yield close to competitive market outcomes, and a merger takes place that significantly reduces the elasticity of the residual demand curve that the combined firm faces, this could cause the combined firm to reduce its fixed-price forward contract obligations in order to have more freedom to withhold output to increase prices in the short-term wholesale market. Thus, in our view a major challenge to finding appropriate remedies for the competitive effects of such a merger requires finding a package of divestitures that causes the merged firm to maintain a high level of fixed-price forward contracts.

Before proceeding with a discussion of how to apply these methods to determine the competitive impacts of a proposed merger and assess the impact of potential divestiture packages in a bid-based wholesale electricity market, there is one more important feature of electricity supply industries that the analysis must address.

\*\*Wolak (2007) demonstrates that forward contracts obligations can commit an expected profit-maximizing supplier to a lower average cost pattern of output within the day, on the order of 5 to 10 percent lower. Therefore, if the firm expects to sell its output at the same price in the forward market or in the short-term market, it may find that signing a fixed-price forward contract that commits it to a lower average cost pattern of daily output is profit-maximizing on an expected value basis.\*\*
Transmission Constraints and Market Definition

Applying the HMG market definition test, we believe that the relevant product market is straightforward to determine because of the limited ability of consumers to substitute away from electricity. The geographic product market definition is complicated by the fact that the ability of consumers to substitute away from local sources of electricity in favor of more distant sources is limited by the capacity of the transmission network, the operating behavior of other generation units, and the level of demand at other locations in the transmission network. In general, when there is no congestion in the transmission network, the relevant geographic market is the entire PJM Interconnection because any attempt by a local supplier to raise the price at one location or set of locations in the transmission network will be undone by substitution to sources of supply at other locations in PJM. Therefore, during all hours of the year without congestion, the relevant product is wholesale electricity, and the relevant geographic market is the entire PJM Interconnection.

Transmission network constraints can significantly reduce the number of generation units and independent suppliers that are able to serve a location or set of locations in the transmission network. Returning to the three-firm example, suppose that (1) firm 3 is distant from either firm 1 or firm 2, and (2) there is a transmission line with finite capacity between firm 3’s location and the location of firms 1 and 2. Figure 1-7 illustrates the impact of this transmission constraint on construction of the residual demand curve facing firm 1, using the same supply curves as in Figure 1-2. The only difference is that a maximum of 300 MWh of firm 3’s supply can actually compete against firm 1 and firm 2 because that is the capacity of the transmission interface between firm 3’s location and the location of firms 1 and 2. This transmission constraint implies that firm 3’s effective supply curve for the purposes of computing firm 1’s residual demand curve becomes vertical at a supply of 300 MWh. Figure 1-7 plots the residual demand curve faced by firm 1 with this transmission constraint taken into account. For all output levels, this curve is at least as steep or steeper than the residual demand calculated in Figure 1-3, which does not account for transmission constraints.

This example demonstrates how transmission constraints can reduce the opportunities for consumers to shift to alternative sources of supply and shrink the geographic size of the market. In Figure 1-7, at prices above $60/MWh (the value where the supply from firm 3 is equal to 300 MWh) firms 1 and 2 no longer face competition from firm 3. In this sense, the transmission network has limited the size of the market in which firms 1 and 2 compete. Therefore, comparing a merger between firms 1 and 2 in a world with infinite transmission capacity between locations in the transmission network with this same merger with finite transmission capacity between locations in the network can yield very different results, depending
on the locations of firms 1 and 2 and the location of competitors to these firms. Both the DOJ complaint and the FERC merger authorization decision emphasized that without appropriate mitigation the merger was likely to harm competition in smaller geographic markets in the PJM Interconnection created by transmission congestion."

MERGER ANALYSIS IN BID-BASED WHOLESALE ELECTRICITY MARKETS

This section describes how generation unit willingness-to-supply curves, market prices and quantities, and firm-level fixed-price forward contract obligations can be used to assess the likely competitive impact of a proposed merger.

Returning to the three-firm market example, let us assume that firms 1 and 2 are the merging parties. Following the procedure outlined above, the premerger residual demand curves for firms 1 and 2 are, respectively, the market demand minus the willingness-to-supply curves of firms 2 and 3 and the market demand minus the willingness-to-supply curves of firms 1 and

3. The residual demand curve of the merged entity is the market demand minus the willingness-to-supply curve of only firm 3.

Figure 1-8 plots the marginal cost curve, MC₂, and D₂, the residual demand curve for firm 2 for the willingness-to-supply curves in Figure 1-2. The best-reply price and quantity pair for firm 2, when it faces residual demand curve D₂, is ($60/MWh, 100 MWh). Figure 1-9 plots the residual demand curve for the merged entity, D₃, and the marginal cost curve for the merged entity, MC₃, which is equal to the horizontal sum of the marginal cost curves of firms 1 and 2. Figure 1-10 computes the intersection of MR₃, the marginal revenue curve of the merged entity, with the marginal cost curve of the merged entity to find the best-reply output and market price for the merged entity with no fixed-price forward contract obligations. This intersection occurs at an output level of 237 MWh for the merged entity, which implies a best-reply price of $72/MWh. Therefore, as a result of the merger, firms 1 and 2 find it unilaterally profit-maximizing to raise prices from $60/MWh to $72/MWh and reduce the amount that they jointly produce from 300 MWh to 237 MWh.

Figure 1-11 performs this same counterfactual merger calculation for D₃, the same value of the residual demand curve for the merged entity as in Figure 1-10, on the assumption that firm 1 has fixed-price forward contract obligations of 100 MWh and that firm 2 has fixed-price forward contract obligations of 50 MWh, so that the merged entity has 150 MWh of fixed-price forward obligations. Figure 1-11 follows the logic presented in Figure 1-6 for computing the best-reply price and quantity with fixed-price obligations.
FIGURE 1-9 Marginal Cost Curve and Residual Demand Curve for Merged Firm

FIGURE 1-10 Calculation of Best-Reply Price and Quantity for Merged Firm with No Fixed-Price Forward Contract Obligations
forward contract obligations for the merged entity. The best-reply quantity for the merged entity is 289 MWh, with a corresponding best-reply price of $62/MWh. This price is significantly lower than the postmerger best-reply price without fixed-price forward contract obligations of $72/MWh and very close to the premerger market price of $60/MWh. This result demonstrates a key factor in the merger analysis: For a high enough level of fixed-price forward contract obligations for the merged entity, the competitive impacts of the merger are very small. By this logic, the merger remedy should provide the strongest possible incentives for the combined entity to maintain a high level of fixed-price forward contract obligations relative to its expected output level.

Comparing the residual demand curves of firms 1 and 2 in Figures 1-4 and 1-8 with the residual demand curve of the merged firm in Figure 1-10, two types of differences emerge. First, the merged entity faces a residual demand at every price level that is larger than the residual demand faced by either firm 1 or firm 2. Second, the merged entity’s residual demand curve is steeper than the one faced by either firm 1 or firm 2. This implies that the price increase that results from a 1 MWh reduction in output by the merged entity is always greater than the price increase that either firm 1 or firm 2 could bring about by a 1 MWh reduction in its output. Both of these factors imply a greater ability of the merged entity to withhold output to raise
prices in the absence of fixed-price forward contract obligations than was true for either party alone before the merger.

Finding the Appropriate Merger Remedy

We believe that there are two important issues raised by our graphical analysis of the competitive impacts of mergers in bid-based wholesale electricity markets. The first is whether the combination will enhance the ability of the merged entity to cause transmission congestion and therefore limit the amount of competition that its generation units face from other suppliers. The second is what level of fixed-price forward contracts will be chosen by the combined entity after the merger.

Modeling Transmission Congestion

The discussion surrounding Figure 1-7 demonstrates that transmission constraints can increase the steepness of the residual demand curves faced by suppliers. Depending on the capacity of the transmission network that connects locations that are served by the merging parties and their competitors, a merger can substantially increase the ability of either party to segment itself from competition by causing transmission congestion. Because the merging parties each owned a substantial amount of generation capacity in PJM East, which can become electrically separated from the remainder of the PJM Interconnection, the merger could increase the frequency and duration of transmission congestion that reduces the amount of competition faced by the merged entity.

When there is transmission congestion between PJM West, where there are many low-variable-cost generation units, and PJM East, where the PECO and PSEG service territories are located, the relevant geographic market for the merger analysis is PJM East. During these hours, suppliers located in PJM West are unable to limit the ability of suppliers in PJM East to raise wholesale prices because transmission constraints prevent any more electricity produced in PJM West from being consumed in the PJM East region.

To assess whether the merger would have increased the opportunities for the combined entity to segment the market, the combined firm’s residual demand curve can be computed under the assumption of transmission congestion that eliminates the ability of certain generation units to compete against it. In this case the residual demand curve that the merged entity faces excludes all generation units owned by nonmerging parties that are located on the other side of a congested transmission interface. For the proposed PSEG and Exelon merger, this requires excluding all suppliers located outside of PJM East from the merged entity’s residual demand curve during hours when there is congestion into PJM East. This reduces the price elasticity of the residual demand curve faced by the merged entity.
during congested hours, which results in higher estimated postmerger prices."

Postmerger Forward Contracting Decisions

The graphical analysis described above demonstrates that the incentive to exercise unilateral market power by a firm with a substantial ability to raise prices can be significantly limited or even eliminated by an appropriate choice of the level of its fixed-price forward contract obligations. A state or federal regulator or antitrust authority is likely to find it impossible to set the quantity of fixed-price forward contract obligations that the merged entity must hold into the indefinite future. Consequently, if these agencies would like the combined entity to maintain a high level of fixed-price forward contract obligations relative to its expected output levels, then a precondition for the merger must be sufficient divestitures of generation capacity to ensure that the merged entity will find it unilaterally profit-maximizing to sign sufficient fixed-price forward contracts into the indefinite future to limit its incentive to exercise unilateral market power in the short-term wholesale market.

State public utility commissions can also mandate (or at least provide very strong financial incentives) for retailers subject to their jurisdiction to maintain high levels of fixed-price forward contracts (relative to their final demand obligations) signed far in advance of the delivery date. In this case, the Pennsylvania Public Utilities Commission’s requirement that PECO, which is owned by Exelon, freeze retail rates until 2010 accomplished that goal. However, we believe that if the wholesale competition concerns associated with the merger were not addressed, retailers purchasing far in advance of delivery may have needed to pay higher prices for these forward contracts.

"PECO and PSEG each owned transmission assets in the PJM East region, so the merger would also have increased the concentration of transmission facility ownership in PJM East. However, we believe that this was unlikely to enhance the ability of the combined entity to use these transmission assets to cause congestion between PJM West and PJM East. Both before and after the merger, the PJM Interconnection would allocate the use of these transmission facilities to the generation units based on their willingness to supply energy. The PJM Interconnection market rules requiring equal access to all transmission facilities in the PJM region imply that the combined entity after the merger, or PECO and PSEG before the merger, could not deny any market participant access to these facilities. Furthermore, it is possible that the combined entity could increase the frequency of PJM East’s becoming a separate market by increasing the frequency and duration of transmission outages. This strategy was very unlikely to be profitable because the merged firm would also own a substantial amount of low-variable-cost generation outside of PJM East that it would like to use to serve its load obligations in PJM East. Page 18 of the merger announcement presentation contains a figure with the estimated variable cost in 2006 for all PSEG and Exelon generation units in PJM. It is also likely that the PJM Interconnection would be able to detect and penalize a significant increase in the frequency and duration of transmission outages on the links between PJM West and PJM East. Consequently, we believe that it was unlikely that the increased concentration in transmission ownership would have affected the extent of unilateral market power exercised by the combined entity in the wholesale electricity market."
contracts because the merged entity controlled such a large fraction of existing generation capacity. Divestitures of generation capacity from the merged entity may be necessary to ensure that it is unable to raise prices in the market for long-term contracts.

The residual demand analysis framework can be easily adapted to address the impact of generation unit divestitures on wholesale prices. To study the effect of the divestiture of a specific generation unit or set of generation units, Figure 1-12 plots the residual demand curve of the merged entity under the assumption that 100 MW of low-cost generation is divested by the merged entity. To simplify the analysis, we assume that this 100 MW is inelastically supplied by the new owner, which effectively shifts inward the residual demand faced by the merged entity by 100 MWh. This postdivestiture residual demand curve is denoted by \( D_{M} \), and the postdivestiture marginal cost curve of the merged entity is denoted by \( MC_{M} \). The postdivestiture best-reply price for the merged entity with no fixed-price forward contract obligations is $65/MWh. This is significantly less than the best-reply price for the merged entity with no fixed-price forward contracts of $72/MWh in Figure 1-10. Different divestiture packages affect both the residual demand curve faced by the merged entity and its marginal cost curve.

Because the merged entity faces a more price-elastic residual demand curve postdivestiture at every level of output, it will have a lesser ability to raise prices through its unilateral actions. This greater elasticity of the residual demand curve implies a greater willingness to sign a given quantity of fixed-price forward contracts because the supplier gives up fewer contracts...
opportunities to exercise unilateral market power in the short-term market by signing these contracts. In addition, as a comparison of Figure 1-4 with Figure 1-6 demonstrates, by signing these fixed-price forward contract obligations the supplier precommits to a higher level of output than it would without these obligations.

According to the logic of our model framework, the process of finding the best possible divestiture package must balance a number of competing concerns. First, the merging parties must be willing to accept the divestiture package. Second, the divestiture package should not increase the geographic concentration of generation ownership, or else the merged entity will have greater opportunities to segment the market by causing transmission congestion. Third, the form of the aggregate marginal cost curve of the merged entity affects its incentive to exercise unilateral market power and its incentive to sign fixed-price forward contracts. This marginal cost curve can be altered by divesting various combinations of generation units from the assets owned by the two parties.

To understand this last point, consider the following stark, but relevant, example: Suppose that the 100 MW of divested generation capacity came from the high-variable-cost generation units owned by firm 2. The marginal cost curves in Figure 1-9 show that firm 1 has almost 150 MW of very low-marginal-cost units, whereas firm 2 has less than 10 MW. Figure 1-13 plots the postdivestiture average marginal cost for the merged entity, assuming that 100 MW of divestitures comes from the high-variable-cost units owned by firm 2 instead of the low-variable-cost units owned by firm 1, as is the case in Figure 1-12. For consistency with Figure 1-12, this 100 MW of

![Figure 1-13 Postdivestiture Best-Reply Price and Quantity for Sale of High-Cost Generation from Firm 2](image)
divested capacity is assumed to be inelastically supplied by the new owner, so the residual demand curve faced by the merged entity is the same as in Figure 1-12. The intersection of postdivestiture marginal cost function $\text{MC}_M$, with the marginal revenue curve associated with the postdivestiture residual demand curve yields a best-reply price of $\$61/\text{MWh}$, which is lower than the best-reply price in Figure 1-12. This analysis illustrates the importance of properly choosing where in the merged entity's marginal cost curve to divest generation capacity in order to limit the adverse impacts of a proposed merger.

This example is relevant to the PSEG and Exelon merger because firm 1 can be thought of as Exelon, and the low-variable-cost units are its nuclear generation facilities. Firm 2 can be thought of as PSEG because it owns a much smaller amount of nuclear MW, as well as a significant amount of higher-variable-cost natural gas-fired generation units. In this example, for the same total MW of generation capacity divested and the same residual demand curve of the merged entity, selling off only high-marginal-cost units reduces the adverse price effects by more than does selling off the same total MWs of capacity of low-marginal-cost units.

This logic also leads to differences in the incentives for the postdivestiture merged entity to enter into fixed-price forward contract obligations. By leaving the merged entity with only low-marginal-cost units, the incentive of the firm to sign fixed-price forward contracts is much higher because the firm knows that it has significantly fewer profitable opportunities to exercise unilateral market power in the short-term market and may face a sustained period of market-clearing prices below its average cost. Conversely, if the merged entity is left with a significant amount of high-marginal-cost units, it will have less of an incentive to sign fixed-price forward contracts because it knows that it is giving up many more opportunities to exercise unilateral market power in the short-term market.

THE PSEG-EXELON MERGER ANALYSIS

This section first describes the initial divestiture proposal made by the parties to address the competition concerns associated with the proposed combination. It then discusses the results of the various merger reviews and the preconditions agreed to by the merging parties and reviewing agencies: the FERC, the Pennsylvania Public Utility Commission, the DOJ, and the NJBPU. This section concludes with our analysis of why the parties eventually decided not to move forward with the merger.

Initial Merger Proposal

In their initial submission to the FERC, the merging parties acknowledged potential competition problems for the PJM Interconnection and PJM East
geographic markets. The FERC (2005, p. 7) notes that, depending on how the Herfindahl-Hirschman Indexes (HHIs) are computed, postmerger HHIs in PJM East ranged from 2057 to 2492, and merger-related changes in HHIs ranged from 848 to 1067, which is well above what would be considered cause for concern in a highly concentrated market. With the initial FERC merger application, the companies proposed to divest a total of 2900 MW of generation capacity, subject to minimums on characteristics of the units sold: approximately 1000 MW of peaking capacity and 1900 MW of mid-merit capacity, with at least 550 MW that was coal-fired. The combined entity proposed to complete this sale as soon as possible: within eighteen months of the close of the merger.

The parties also proposed a "virtual divestiture" of 2600 MW base-load nuclear generation units, which included 2400 MW in PJM East. This virtual divestiture would take one of two forms: (1) fixed-price long-term contracts for at least fifteen years or the life of the generation unit or (2) an annual auction in twenty-five MW blocks of three-year firm entitlements at fixed prices to the output of the nuclear generation units. This virtual divestiture was designed by the merging parties to limit the incentive of the combined entity to increase prices in the short-term market without losing the benefits that were associated with applying Exelon's nuclear generation plant operating expertise to all of the nuclear facilities owned by the combined entity.

The state regulatory filings by the parties addressed the state-level public benefits requirements associated with the merger. In these filings, the merging parties emphasized that the merger would enhance the ability of individual companies to provide cost-effective, safe, and reliable service without any price increases to the retail customers of PECO, ComEd, and PSEG. These filings emphasized the existence of operating cost savings from the merger associated with scale, scope, and best-practice sharing, without much detail as to the sources of these benefits except for the discussion of Exelon’s nuclear plant operation expertise’s being applied to PSEG’s facilities. The filings also noted that the combination would result in a reduction of the combined entity’s workforce of approximately 5 percent and that, to the maximum extent possible, these workforce reductions would occur through attrition, although severance programs could also be utilized.

"Generation capacity is often characterized by when during the day a unit is expected to operate. Base-load units typically operate during all hours of the day. Mid-merit or intermediate units operate during the vast majority of hours of the day. Peaking capacity operates only during the highest demand (peak) hours of the day.


*Pages 25 of the merger announcement presentation (http://media.corporate-ir.net/media_files/Irel/12124290/pdf/EXC_PSEG_AnalystPres_122004.pdf) discusses these sources of cost savings from the merger.
The parties claimed annual pretax merger synergies in the first year of $400 million, with this number increasing to $500 million for the second year. Approximately 86 percent would come from operating cost savings from eliminating redundant activities at the two companies and realizing economies of scale in the acquisition of inputs. According to the joint proxy statement and prospectus for the 2005 annual meeting of Exelon shareholders, the remaining 14 percent would come from increased capacity utilization at PSEG’s nuclear facilities. This joint proxy statement estimated the cost to achieve these synergies to be $450 million in the first year following the completion of the merger and approximately $700 million over the four years following the merger.

Based on these figures, the joint proxy statement estimated the net benefits of the merger to be about $200 million, which did not leave a substantial amount of merger benefits to share with consumers in Pennsylvania and New Jersey to meet the state-level public benefit test for merger approval in both states—an issue very relevant to the final outcome of the proposed merger. We believe that the prospect that the nuclear operating cost savings and increased efficiency for PSEG’s nuclear facilities would be realized without the proposed acquisition (because of the separate nuclear operating agreement between Exelon and PSEG noted earlier) is likely to have further reduced the benefits that the parties directly attributed to the merger. The fact that the parties had almost two years of experience with the joint operating agreement at the time that the negotiations with the NJBPU broke down is also likely to have reduced their desire to share the nuclear plant operating cost savings and revenue increases with consumers in New Jersey.

The state-level filings also emphasized that PECO, ComEd, and PSEG would remain as separate corporations after the merger with headquarters in Philadelphia, Chicago, and Newark, respectively. Finally, the filings emphasized that local charitable contributions and support for economic development within each state would continue at the same or higher levels after the merger.

The Merger Approval Process and Modifications of Divestiture Packages

The FERC merger review process identified several conceptual and factual errors in the initial analyses filed by the merging parties to justify their initial divestiture packages. A number of parties also disputed proposals by the merging parties to restrict the set of potential buyers of the divested

---


Page 61
assets to entities without large preexisting market shares in the PJM interconnection. In order to respond to these complaints and in order to relax the restrictions on the preexisting market shares of potential buyers of the divested assets, in May 2005 the merging parties upped the initial 2900 MW divestiture proposal to 4000 MW, composed of roughly 700 MW of base-load units, 2100 MW of mid-merit units, and 1200 MW of peaking units. On July 1, 2005, the FERC approved the merger with this divestiture package and the 2600 MW virtual divestiture of nuclear capacity described above.

The next to act was the Pennsylvania Public Utility Commission, which approved the merger with the proposed divestiture package in late January 2006. As part of the merger agreement, PECO agreed not to increase retail prices until 2010. It also agreed to increase the amount of energy usage per household that could qualify for a reduced price of electricity and to spend $1.2 million on additional consumer outreach to acquaint low-income consumers with this program. PECO also pledged to maintain its corporate headquarters in Philadelphia through at least 2010. Subject to these terms and conditions, the Pennsylvania Public Utility Commission found that the parties had met the public interest standard for approval of the merger.

From the early summer of 2005 until June 2006, the DOJ undertook a comprehensive analysis of the merger. In late June 2006 the DOJ reached a settlement with the merging parties that involved the divestiture of 5600 MW of fossil-fuel capacity. This settlement specified the divestiture of all of the units at six generation facilities in the PECO and PSE&G service territories. All of the facilities in this divestiture package had variable costs that were close to the average market clearing price in PJM during a number of hours of the day during the sample period, which is consistent with the pattern of divestiture recommended by the logic in Figures 1-12 and 1-13: For a given MW quantity of generation divestitures, selling units with variable costs in the range of actual market clearing prices is likely to result in lower postdivestiture prices than is selling the same amount of capacity in low-variable cost units.

It is also important to note that the plant divestitures required by the DOJ would have resulted in the merged entity’s owning less fossil fuel-generation capacity in the PJM East region than PSEG owned before the merger, which is consistent with the DOJ’s desire to address the PJM East competition issues noted in its complaint. Under the terms of the DOJ

“Page 18 of the merger announcement presentation (http://media.corporate-ir.net/media_files/doc/12124256/pdfs/EXC_PSEG_AnalystPres_122004.pdf) shows the estimated variable cost in 2006 for all PSEG and Exelon generation units in PJM.”
agreement, the merged entity would have had to enter into an agreement to sell this capacity within 150 days after the merger closed. In addition, for a period of ten years following the merger, the combined company would have had to obtain the DOJ’s prior approval before acquiring or obtaining control of power plants in the PJM East region.

No nuclear generation capacity was required to be included in the DOJ’s divestiture package. This is consistent with the analysis in Figures 1-12 and 1-13 described above. An expert witness for the merging parties argued that the combined entity would not withhold nuclear capacity to raise wholesale prices because the extremely low variable cost of this capacity relative to the typical market price in the PJM Interconnection makes withholding output from these units very costly. As noted earlier, a major source of benefits claimed by the parties from the proposed merger would be the increased efficiency at PSEG’s nuclear facilities to be realized by implementing Exelon’s management practices, and the incentive to realize these operating cost savings and revenue increases would be greater if Exelon owned rather than simply operated PSEG’s nuclear assets for a finite period of time, which would be the case for any of the virtual divestiture packages. This logic and the analysis in Figures 1-12 and 1-13 suggest that the DOJ was wise to focus its efforts on fossil fuel units with variable costs at or near average hourly prices in PJM.

The only outstanding merger review process at this point was that by NJBPU, the New Jersey public utility regulatory body. The NJBPU issued an order on June 20, 2005, requiring Exelon and PSEG to prove that PSE&G customers and the state of New Jersey would benefit from the merger and that the merger would not lead to adverse effects on competition, employees of PSE&G, and the reliability of electricity supply to the state. Negotiations between the merging parties and the NJBPU proceeded for almost three months after the agreement with the DOJ was announced. In its Form 10-Q to the Securities and Exchange Commission for the quarter ending June 30, 2006, PSEG noted that the merging parties recently had made a materially enhanced cash settlement offer to the NJBPU, which they believed would provide substantial positive benefits to customers and to the state of New Jersey. This cash settlement offer could have been used for a variety of customer and state benefits, primarily price reductions for PSE&G’s customers. The Form 10-Q report noted that PSEG’s earnings and cash flow would be materially reduced in the near term as a result of this settlement offer. In this document, PSEG also expressed the hope that a settlement with the NJBPU could occur in time to allow the merger to close by the end of the third quarter of 2006.


This document is available at http://yahoo.brand.edgar-online.com/EFX.dll/EDGARpro.dll?FetcherFilingHTML1?SessionID=n6NXnL0Q_Sq7Fjg&ID=4566197.
In mid-September 2006 PSEG and Exelon announced that Exelon had given PSEG formal notice of the termination of the merger agreement. At the same time, the parties withdrew their application for approval of the merger at the NJBPU, which had been pending for more than nineteen months. Although neither party made any statements about the specific reasons for the termination of the merger, the major points of difference were the magnitude of benefits that New Jersey ratepayers could expect to see from the merger and how the wholesale market competitive effects associated with the combination would be addressed. Based on the public disclosures by PSEG in its Form 10-Q report, it seems likely that the amount of the public benefits settlement desired by the NJBPU may have been sufficiently large relative to the other concessions to which the merging parties had agreed with the other reviewing agencies to cause the net benefits that PSEG and Exelon expected to realize from the merger to be very close to zero or even negative. The estimated merger net benefits of $200 million referred to earlier provide further credibility to this explanation for the termination of the merger agreement.

LESSONS FOR PARTICIPANTS IN FUTURE MERGERS IN WHOLESALE ELECTRICITY MARKETS

The quantitative methods presented here can be applied to a merger in any bid-based wholesale electricity market. As the number of these electricity markets in the United States increases, the opportunities to apply these methods are likely to increase. These methods allow a detailed quantitative assessment of the competitive impacts of proposed packages of generation unit divestitures that would address the incentives for the merged entity to maintain a high level of fixed-price forward contract obligations following the merger and limit the opportunities for the merged entity to segment a larger geographic market from the remainder of that market in order to raise the wholesale prices paid to its generation units.

The ultimate outcome of this proposed merger suggests that few mergers involving generation unit owners in wholesale electricity markets will be able to survive the multistage, federal and state antitrust and regulatory approval process and still provide value to the shareholders of the merged companies. The public benefit standard applied by most public utility commissions provides state governments with a substantial ability to extract financial concessions from the merging parties that may cause the merging parties to terminate potentially beneficial mergers.

It is unclear whether the proposed merger was an example of this phenomenon. The description of the merger demonstrates that it was difficult to find tangible and sizeable benefits that could be solely attributed to the merger. However, it is an open question whether the merger was in the broad public interest in the sense that it benefited all consumers served by
the PJM Interconnection, and the merger was terminated because the New Jersey and Pennsylvania public utility commissions demanded too many explicit financial concessions for ratepayers and citizens. Exelon’s upper management appears to have felt that this was the case.” In late May 2006, before the DOJ settlement was announced, a major investor in Exelon, the hedge fund Duquesne Capital, asked the board of directors of Exelon to terminate the merger agreement, stating that it was “a deal which has gone very bad with the passage of time.”

The public benefit criteria for merger analysis traditionally used for mergers involving public utilities within the boundaries of states may therefore be inappropriate in those instances when the merging utilities participate in multistate wholesale markets such as PJM. This local public benefit view of mergers of electric utilities involved in multistate wholesale markets may need to be revised before these sorts of mergers can be approved.

REFERENCES


See “NJBPU Staff’s Demand for Shared Savings Is Biggest Obstacle to Exelon/PSEG Merger,” Electric Utility Week, June 5, 2006, p. 4, for a statement by Exelon upper management concerning the major obstacles to completing the merger at that time.


Oracle’s Acquisition of PeopleSoft:  

R. Preston McAfee,  
David S. Sibley, and  
Michael A. Williams*

**INTRODUCTION**

On June 6, 2003, the Oracle Corporation made an unsolicited cash tender offer for all of the outstanding shares of PeopleSoft, Inc. Oracle and PeopleSoft are enterprise software companies that develop, manufacture, market, distribute, and service software products that are designed to help businesses manage their operations. Together with SAP AG, they are the three largest companies in the industry. Oracle’s total revenues in fiscal year 2004 were $10.1 billion, while PeopleSoft and SAP AG had total revenues in 2003 of $2.3 billion and $8.0 billion, respectively. As discussed in detail below, all three firms produce enterprise resource planning (ERP) software that enables companies to operate their human resources, finances, supply chains, and customer relations.

In February 2004, the U.S. Department of Justice (DOJ), together with the states of Connecticut, Hawaii, Maryland, Massachusetts, Michigan, Minnesota, New York, North Dakota, Ohio, and Texas (plaintiffs) filed suit

*Preston McAfee served as an economic expert on behalf of the U.S. Department of Justice, Antitrust Division, in the matter of _U.S. et al. v. Oracle_. Michael Williams and the ERS Group were retained by the U.S. Department of Justice’s Antitrust Division to assist in the development of the economic analysis underlying McAfee’s testimony. During this period, David Sibley was deputy assistant attorney general for economics at the Antitrust Division.*
to enjoin permanently Oracle’s acquisition of PeopleSoft. The case raises several interesting antitrust issues. The DOJ’s case was founded on the theory that the merger would adversely affect ERP buyers because of unilateral competitive effects. That is, the DOJ did not assert that the merger would lead to tacit or explicit collusion but rather that the merger would cause the merging parties to cease competing, which would adversely affect customers for whom the two companies’ products were the first and second choices (holding constant rivals’ competitive strategies). In particular, the DOJ argued that buyers procured ERP software in a manner appropriately modeled by auction theory and that the proposed merger would eliminate PeopleSoft as a bidder. As part of its analysis, one of the DOJ’s experts presented a merger simulation model based on auction theory. United States v. Oracle represents the first case in which a merger simulation was used in court. What should be the role of merger simulations in analyzing the competitive effects of horizontal mergers? In particular, what should be the role of market definition in merger simulations of unilateral effects cases?

The DOJ and Oracle also relied extensively on customer testimony in providing evidence pertaining to market definition and the likely competitive effects of the proposed merger. What should be the role of customer testimony in analyzing these two issues?

After a trial in U.S. District Court for the Northern District of California, Judge Vaughn Walker ruled on September 9, 2004, that the DOJ had not proven its case that the proposed merger would violate U.S. antitrust law. The DOJ announced on October 1, 2004, that it would not appeal Judge Walker’s decision. In December 2004, PeopleSoft’s board of directors accepted Oracle’s $10.3 billion offer.

MARKET BACKGROUND: ENTERPRISE APPLICATION SOFTWARE

The proposed merger of Oracle and PeopleSoft raised concerns about software products belonging to the broad category commonly called “enterprise application software” (EAS), of which ERP is one type. These products are used to automate the performance of necessary business functions. Important segments (or “pillars”) include the following: (1) “human resources management” (HRM) software, which automates payroll services,
recruiting, training, and benefits administration; (2) "financial management systems" (FMS) software, which automates the general ledger, accounts payable and receivable, and asset management; (3) "supply chain management" (SCM) software, which assists in the control of inventory, manufacturing, and distribution; and (4) "customer relations management" (CRM) software, which manages the entire life cycle of a sale, from the development of customer prospects to customer support and service. Each pillar may contain from thirty to seventy modules. Some companies sell suites containing modules from more than one pillar. Such combinations of pillars are referred to as "ERP suites." Typically, an ERP suite is a collection of packaged software that integrates most of a firm's data across most of its activities. When an individual pillar is sold on a stand-alone basis, it is known as a "point solution" or "best-of-breed solution."

Businesses vary greatly in size, complexity, and the efficiency of the embedded or legacy information technology infrastructure used to support their operations. As a result, businesses vary greatly with respect to the features that they value in ERP and their willingness to pay for these features. A product that meets the requirements of one large, multinational corporation with global operations may not meet the requirements of another large, multinational corporation or the requirements of a small, single-establishment business. For these reasons, ERP software exhibits considerable product differentiation, with vendors often focusing on the requirements of specific industries (or "verticals") such as banking, health care, and government. Vendors also differ with respect to the types of software that they develop, with some firms focusing on off-the-shelf products that serve the requirements of smaller firms with relatively simple business operations, and with others focusing on complex, customizable software and ancillary services designed to meet the specific requirements of large, complex enterprises (LCEs). For LCEs, the fees to license and maintain ERP software are often only 10 percent to 15 percent of the total cost of ownership, which also includes the costs of personnel training, consulting, and integrating the new program with the customer’s legacy software and databases.

Large firms that purchase complex, customizable software typically rely on competitive bids to procure solutions to their business requirements. Such firms identify a relatively small number of vendors that are capable of meeting their requirements, send them requests for proposals (RFPs), and engage in protracted negotiations with the vendor(s) that submit the most attractive initial responses.

"Customizable software" means that the software can be configured to fit a customer's requirements. This does not mean that the fundamental software code is modified for a given customer. The software that Oracle, for example, sells to a large, multinational company is identical to the software it sells to a small company. "Customization" is achieved through the use of software settings that can be configured to match a customer's requirements.
THE DOJ’S Case

Competitive Effects Theory

The DOJ began by distinguishing three categories of EAS: (1) “off-the-shelf PC-based products suitable for many small businesses; (2) relatively inexpensive software with limited capability that must be professionally installed and maintained and is suitable for "mid-market" firms; and (3) "high-function enterprise software” required by "enterprise customers" or LCEs. High-function products support thousands of simultaneous users and tens of thousands of simultaneous transactions, integrate seamlessly across pillars (specifically, HRM and FMS modules), and are sufficiently flexible to support the unique business processes of each LCE.

Critical features of high-function enterprise software include the ability to support business operations that span (1) multiple jurisdictions with multiple currencies and languages; (2) multiple legal entities or divisions within the business; and (3) multiple lines of business. High-function enterprise software is also distinguished by (1) its total cost of ownership, which amounts to millions of dollars; (2) the relatively long period (e.g., several months) taken by customers to make a purchase decision; and (3) the difficulty in implementing the software. Enterprise customers will not consider vendors that cannot provide continuous technical support and continuous enhancements to the product’s capabilities over its long life. LCEs are also unwilling to consider vendors that lack a track record of successful implementation of high-function enterprise software suites.

The typical procurement cycle of large, complex enterprises includes the following steps: The enterprise analyzes and identifies its requirements; determines the expected return on investment and prepares a budget; forms a selection committee; establishes detailed functional requirements; issues a request for information (RFI) to prescreen possible vendors; issues requests for proposals (RFPs) to qualified vendors; arranges for demonstrations from three to five qualified vendors; receives bids; negotiates price and terms with two to three vendors; and then selects a final vendor.

Based on this description of the products and how they are procured, the DOJ focused on high-function products and argued that the primary rivals were Oracle, PeopleSoft, and SAP AG’s U.S. subsidiary, SAP America Inc. Plaintiffs, thus, viewed the case as a "3-to-2" merger in which:

Unlike consumer products with uniform pricing, competition in this case involves a bidding process that is separate for each customer. Copious evidence documents the fact that discounts vary considerably across

*Plaintiffs’ Post-Trial Brief, p. 1.*
customers, depending on the particular circumstances of each customer and the competition to supply each customer. Systematic analysis of Oracle’s data on E-Business Suite sales and from its discount approval forms and sales representative survey reports indicates that Oracle discounts significantly more than otherwise when in competition with PeopleSoft. Because the price competition to sell high function HRM and FMS software is specific to the particular customer, the effects of the merger differ across customers according to the significance of the head-to-head competition between Oracle and PeopleSoft. Based on this analysis, the DOJ determined that buyers procured ERP software in a manner appropriately modeled by auction theory. Since the merger would eliminate PeopleSoft as an independent bidder, the DOJ argued that Oracle would be able to win postmerger procurements with higher bids.

The DOJ did not present evidence at trial regarding possible adverse welfare effects resulting from coordinated effects. However, in its post-trial brief, the DOJ briefly addressed the likelihood that the proposed merger would lead to increased prices via coordinated effects. The DOJ observed that coordination was likely, given the high concentration in the relevant markets, and noted further that concentration in winner-take-all competitions for individual customers could take the form of an allocation of customers.

**Market Definition**

The DOJ applied standard principles of market delineation as set forth in the *Horizontal Merger Guidelines*. In particular, the DOJ defined two relevant product markets: high-function HRM and high-function FMS enterprise software. Preconfigured versions of these software modules sold to mid-market firms were excluded from the relevant market on the grounds that these versions would not meet the requirements of LCEs and thus would not constrain the prices charged to these customers. The relevant geographic market was defined to be the United States, based on the DOJ’s claim that a hypothetical monopolist consisting of all U.S. sellers of high-function HRM or high-function FMS would find it profitable to increase prices by a small but significant and nontransitory amount. The DOJ also argued that sellers engaged in price discrimination in the relevant markets. However, the DOJ did not use the price discrimination approach described in the *Horizontal Merger Guidelines* to define the relevant markets.

*Plaintiffs’ Post-Trial Brief, p. 33.*

‘Price discrimination is the practice of charging different prices for similar products when the price differences do not reflect underlying differences in cost.'
In support of this market definition, the DOJ’s case offered several strands of evidence. Reports by independent market research firms were cited in support of the view that mid-market software is not a good substitute for high-function software and in support of the DOJ’s conclusions regarding market participants. The “Big 5” consulting firms are often used by LCEs to assist in structuring and managing the procurement process for EAS. The DOJ pointed to two senior executives of Big 5 firms who supported the DOJ’s market definition and the list of market participants. Evidence from Oracle, including its descriptions of the industry, its customer surveys, and its plans to gain share in the “mid-market,” were also cited in support of the DOJ’s market definition. Evidence from PeopleSoft (a market participant) and from Microsoft (alleged to be a new entrant) was also cited in support of the view that mid-market software was unsuitable for many LCEs. Evidence from PeopleSoft’s acquisition of JD Edwards, a producer of mid-market software, was also offered in support of the view that mid-market software was not a substitute for the products required by LCEs. Finally, the DOJ relied on the results of a merger simulation model that found that the proposed merger would result in price increases of 13 percent to 30 percent for high-function HRM software and 5 percent to 11 percent for FMS software.

The DOJ characterized these market definitions as appropriate for identifying the main economic forces that constrain prices. The plaintiff concluded that differences in the products of Oracle, PeopleSoft, and SAP America were much smaller than the differences between these firms’ products and those of other software vendors. The plaintiff considered whether mid-market software was a close substitute for high-function HRM and FMS products and reached the conclusion that this software was outside the relevant product market. The DOJ also concluded that self-supply, legacy systems, outsourcing, Microsoft’s new products, freeware, and “point solutions” were not sufficiently close substitutes to be included in the market.

The plaintiff also analyzed “Discount Authorization Forms” produced by Oracle and found that competition from PeopleSoft was identified 122 times as the justification offered by Oracle salespeople in their requests for authorization to provide selected discounts. SAP America, the next-most-cited firm, was listed only 81 times. The DOJ concluded that these data supported a product market definition limited to high-function FMS and HRM software sold to LCEs.

The Big 5 consulting firms serving enterprise software customers are Accenture, BearingPoint, Cap Gemini, Ernst & Young, and IBM Global Services.

The simulation model was based on auction theory. It was developed by one of this chapter’s authors (McAfee), who was retained by the DOJ to analyze the likely competitive effects of the proposed merger.
Market Structure

The plaintiff assigned market shares using sales data produced by Oracle, PeopleSoft, and third parties. Transactions below $500,000 were excluded on the grounds that they were likely to be for mid-market products outside the relevant market. For FMS software, total sales amounted to $114 million in 2003. PeopleSoft had a share of 32 percent, and Oracle had a share of 17 percent. SAP America, the only large rival, had a share of 39 percent. In addition, AMS had a market share of 10 percent, Microsoft a share of 2 percent, and SCT a share of 1 percent. The merger increased the Herfindahl-Hirschman Index (HHI) by 1064 points, resulting in a postmerger HHI of 3994.

For HRM software, total sales in 2003 were $129 million. Market shares were 51 percent for PeopleSoft and 18 percent for Oracle, with an additional 29 percent for SAP America, 2 percent for SSA, 1 percent for Lawson, and 0.4 percent for SCT. The merger increased the HHI by 1802 points, resulting in a postmerger HHI of 5497.

Accordingly, the DOJ argued that the market was highly concentrated, with three large firms (Oracle, PeopleSoft, and SAP America) and a small competitive fringe; in essence, the proposed transaction was a 3-to-2 merger. The fringe firms were characterized as producing differentiated products that appealed to few customers, with no competitive influence beyond these customers.

Evidence on Competitive Effects

The DOJ focused on the likely unilateral effects of the proposed merger. The process of matching a LCE’s unique requirements with specific features of the vendors’ software could result in unique transactions with unique prices. An analysis of confidential Oracle business records indicated that vendors gained a great deal of relevant information during the competition to serve a customer. Similarly, an analysis of Oracle’s business records showed that prices for the same software sold at the same time depended on factors other than cost, thus establishing the ability of vendors to price discriminate.

To quantify the anticompetitive effects of the merger, the plaintiff undertook three independent analyses using three different data sources. The three approaches taken were (1) a statistical analysis of transactions databases maintained by the merging parties; (2) price regressions; and (3) a simulation model based on auction theory.

A Statistical Analysis of Transactions Data

The transactions databases showed that Oracle competed more frequently with PeopleSoft and with SAP America in the larger deals, that is, the deals more likely to include the high-function software at issue. Also, Oracle was
found to win less often when PeopleSoft was a competitor than when it was not a competitor. Similarly, PeopleSoft was found to win more often when Oracle was not a competitor. The results suggested a localized product space in which Oracle and PeopleSoft were each other’s closest competitors.

**Price Regressions**

The DOJ performed a regression analysis of the price discounts offered by Oracle for its “E-Business Suite.” The regression showed how Oracle’s discounts varied depending on the presence or absence in a given procurement competition of PeopleSoft, JD Edwards, SAP America, Siebel, and niche producers. When PeopleSoft (or any other vendor) was present in a procurement competition, a binary (or “dummy”) variable was set equal to one, and when it was absent, the dummy variable was set equal to zero. The regression was run for thirty-seven transactions with sales greater than $500,000 (see Table 2-1). The coefficient for the PeopleSoft

<table>
<thead>
<tr>
<th>TABLE 2-1</th>
<th>Dependent Variable: Percentage Price Discount on Oracle’s E-Business Suite</th>
</tr>
</thead>
<tbody>
<tr>
<td>PeopleSoft</td>
<td>0.097 (0.049) [0.056]</td>
</tr>
<tr>
<td>JD Edwards</td>
<td>0.071 (0.058) [0.228]</td>
</tr>
<tr>
<td>SAP America</td>
<td>0.097 (0.074) [0.197]</td>
</tr>
<tr>
<td>Siebel</td>
<td>0.030 (0.051) [0.561]</td>
</tr>
<tr>
<td>Niche</td>
<td>-0.052 (0.044) [0.248]</td>
</tr>
<tr>
<td>$500k &lt; Deal &lt;$1M</td>
<td>-0.015 (0.043) [0.723]</td>
</tr>
<tr>
<td>Constant</td>
<td>0.625 (0.040) [&lt;.001]</td>
</tr>
</tbody>
</table>

R² | 0.287 |
N | 37 |

(standard errors in parentheses) [p-values in square brackets]
variable was 0.097. This coefficient means that Oracle offered a 9.7 percentage point greater discount in procurements in which PeopleSoft competed than Oracle’s regression-weighted average discount in procurements in which PeopleSoft did not compete. The regression implies that when there are no competitors to Oracle (so that each dummy variable equals zero), Oracle’s average discount off its list prices equals 62.5 percent, that is, the constant term in the regression. When PeopleSoft is the only additional competitor, Oracle’s discount deepens on average by 9.7 percentage points (the coefficient on the dummy variable for PeopleSoft) to 72.2 percent. This interpretation of the regression is consistent with Judge Walker’s statement that “when Oracle competes against PeopleSoft for the sale of Oracle’s E-Business Suite, the consumer obtains a 9.7 [percentage point] greater discount than when Oracle competes against no one in selling the suite.”

However, the competitive effect of PeopleSoft’s presence goes beyond this special case. The regression shows that PeopleSoft’s presence increased Oracle’s discounts by 9.7 percentage points on average compared with all procurements in which PeopleSoft did not compete. For example, in procurements in which Oracle competed against SAP America alone, Oracle increased its discount by a smaller amount than when it competed against both SAP America and PeopleSoft. In the latter case, Oracle’s total discount equals the coefficient on the PeopleSoft variable (9.7 percentage points) plus the coefficient on the SAP America variable (9.7 percentage points) plus the coefficient on the constant term (62.5 percentage points), for a total discount of 81.9 percent. Thus, the regression is not limited to quantifying only the additional discount offered by Oracle when it competes against PeopleSoft compared with the discount Oracle offers when it competes against no one in selling the suite.”

Merger Simulation Model

The price effects of a merger can be measured quantitatively through a merger simulation (Werden 2005). A simulation uses data from the market—for example, prices and quantities—to calibrate an economic model of competitor interaction. Once calibrated, the model can give quantitative estimates of important economic parameters; in particular, a properly developed model yields parameters that describe the level of

*United States v. Oracle, p. 1169.

“The more general interpretation of the regression results appears to have been missed in subsequent comments on the Oracle case. Coleman (2005) incorrectly concludes that “one of the DOJ’s expert witnesses, provided estimates of the differences in Oracle’s discounts when PeopleSoft was present versus when no competitors were present and showed that discounts were higher with PeopleSoft than with no competitors. . . . [This finding does not provide information as to whether the competition between PeopleSoft and Oracle was unique because it does not show that discounts were higher than when other competitors were present.”
prices in the market depending on the number of competitors. One can then simulate the price effects of a merger by removing one competitor and then predicting postmerger price levels using the original parameters from the model.

Because market data are used to estimate an underlying economic model, the estimated effects of a merger depend upon the model chosen to represent the market. Although there are a number of alternative models, three canonical economic models are primarily used in merger simulations: Bertrand, Cournot, and auctions. Each model uses a different set of rules that describes the choices available to firms. In a Bertrand model, typically used for markets in which firms sell differentiated products, competitors simultaneously choose prices such that those prices maximize their profits, given the prices chosen by all other firms. A Bertrand equilibrium exists when no firm can increase its profits by changing its price.

In a Cournot model, typically used for markets in which firms sell homogeneous products, firms simultaneously choose quantities such that those quantities maximize their profits, given the quantity choices of all other firms. A Cournot equilibrium exists when no firm can increase its profits by changing its output. There are several types of auctions. In a merger context, auctions generally take the form of a procurement auction in which firms compete by offering to supply a product, with the lowest bid winning the auction.

In a Bertrand model, a firm's price depends on the demand elasticity that firm faces. The more inelastic the demand, the higher will be its markup. A merger simulation analyzes consumers' substitution patterns between firms to predict the demand elasticity that the combined firm would face and hence to predict its postmerger prices. With Cournot, a firm's price is a function of the market demand elasticity and the firm's market share, with a higher market share implying a higher price. A simulation based on the Cournot model estimates the market elasticity of demand and uses it with postmerger market shares to predict prices. Finally, with an auction model, while the prices in some particular auctions may not change as a result of a merger (which can occur when the merging firms' products are not the buyer's first and second choices), the reduced number of suppliers can increase the winning bid. As described below, an auction model was used in the merger simulations in the Oracle case.

Merger simulation uses standard economic tools and, properly applied, is firmly grounded on the facts of the relevant market. As such, it avoids problems associated with descriptive analyses based on an expert's intuition. It also can avoid the often difficult issue of market delineation.

\[ \text{price-cost margin} = \frac{\text{price} - \text{marginal cost}}{\text{price}} = \frac{1}{\text{demand elasticity}} \]

\[ \text{price-cost margin with Cournot} = \frac{\text{firm's market share}}{\text{market elasticity of demand}} \]

\[ \text{specifically, a firm's} \ "\text{price-cost margin}" \ (\text{price minus marginal cost, all divided by price}) \ \text{equals the inverse of the demand elasticity for its product.} \]

\[ \text{with Cournot, the price-cost margin equals the firm's market share divided by the market elasticity of demand.} \]
required by a traditional merger analysis (see discussion below). While a merger simulation requires decisions about which products to include, the relevant elasticities will still account for any competitive effects associated with products excluded from the analysis. Merger simulation, however, is dependent on the appropriateness of the underlying economic model. Indeed, since no model perfectly describes a market, picking the correct model involves evaluating the salient features of the market, matching them to a model, and then testing the model against the data.

The merger simulation model used by the DOJ was based on an English auction (i.e., an ascending-price, open-outcry auction) model because that model allows for the presence of multiple bidders and multiple rounds. The analysis used a complete information model in which each software vendor knew the value that the buyer placed on each vendor’s product. In this model, a buyer’s valuation of the \( i \) competitor’s software product was assumed to be \( V_i > 0 \), and \( V \) was assumed to be uniformly distributed over \([0,1]\). Given \( a \) for all competitors, the relative probability that each competitor will win can be calculated. Alternatively, given market shares, the relative values of the \( a \) can be calculated.

The merger simulation used the market shares mentioned above to calculate the relative values of the \( a \). The simulation depended in part on an assumption regarding how much of the value of a software product accrues to the buyer and how much accrues to the seller in the form of the sales price to obtain the absolute values of the \( a \). The DOJ used a range of values for this parameter, from 50 percent of the value accruing to the buyer to 90 percent of the value accruing to the buyer. The merger simulation based on the estimated parameters showed that the price of FMS software in deals greater than or equal to $500,000 likely would increase by 5 percent (using the 50 percent accrual assumption) to 11 percent (using the 90 percent accrual assumption). According to the simulation, the price of HRM software in such deals likely would increase by 13 percent (using the 50 percent accrual assumption) to 28 percent (using the 90 percent accrual assumption).

**ORACLE’S ARGUMENTS**

**Market Definition**

Oracle’s defense identified demand and supply factors that constrained the prices of the products at issue. Oracle argued that the markets defined by the DOJ were too narrow in both their geographic and product dimensions. For geographic markets, the defendant applied the Elzinga-Hogarty (1973) test to PeopleSoft’s data on sales of FMS and HRM software to non-North American customers and to SAP America’s sales of the relevant products to North American customers. The test is based on the physical flow of products between regions, using two measures of a region’s openness: LIFO...
(little in from outside) and LOFI (little out from inside), which are defined as follows:

\[
UFO = \frac{\text{in from outside}}{\text{total supply}} \quad \text{all local consumption}
\]

\[
LOFI = \frac{\text{out from inside}}{\text{total supply}} \quad \text{all local supply}
\]

Elzinga and Hogarty suggest that a given region constitutes a geographic market if both ratios exceed 75 percent or if the average of the two ratios exceeds 90 percent. Based on this test and other evidence, Oracle argued for a worldwide market. Oracle also cited a European Commission decision regarding the proposed acquisition that found the relevant geographic markets for high-function FMS and HMR were worldwide (European Commission 2004, para. 179).

Oracle criticized the DOJ’s product market definition for lacking precision and for appearing to work backward with the aim of limiting the market to customers of Oracle, PeopleSoft, and SAP America. The defendant claimed that many of these customers were not similarly situated and that the DOJ’s market definition did not account for important forces that constrained price. Among these forces were incumbent or legacy systems, outsourcing, and other firms such as AMS and Lawson that had been excluded by the DOJ. Oracle provided examples of LCEs that utilized each of these alternatives.

The defendant also took issue with the DOJ claim that sellers in the relevant markets engaged in price discrimination. Oracle argued that buyers were large, sophisticated customers who would not let vendors gain the critical information necessary to implement profitable price discrimination. In support of these arguments, Oracle’s expert cited research showing that a monopolist must correctly guess customers’ willingness to pay more than 90 percent of the time in order for price discrimination to be profitable (Hausman, Leonard, and Vellturo 1996).

Oracle did not attempt to define the relevant product markets, instead arguing that the DOJ’s definition was incorrect and that the properly defined product market would have been broader than the market defined by the DOJ.

Theory and Evidence on Competitive Effects

Oracle observed that, in a large majority of cases, customers chose to limit the number of competitors in the final round of the procurement cycle to one or two firms and that thus the perceived lack of competition (even before the merger) was only apparent. For the purpose of negotiating the final contracts, customers possessed sufficient buyer power to ensure that
they obtained lower prices, and additional negotiations with a second or third vendor might not provide any additional benefit.

Accordingly, bargaining theory, rather than auction theory, provided the appropriate framework for analyzing likely competitive effects of the proposed merger. Bargaining theory is based on the work of Nobel Prize winner John Nash (1950 and 1953), who showed that under reasonable conditions two parties with equal bargaining power will negotiate an outcome in which each player gets half the total incremental gains from cooperation to both players. This outcome maximizes the product of the gains accruing to the two parties from their deal.

Oracle asserted that the negotiations in question did not resemble English auctions in any important manner. In a large majority of the deals, buyers reduced the number of potential suppliers to one or two. Because an auctioneer would not so limit the number of bidders, Oracle concluded that the buyers used their bargaining power and their value as reference customers, rather than an auction, to obtain low prices. Oracle noted that variation in prices, terms, and conditions among customers may reflect differences in negotiating skills and other related factors (such as whether the negotiation occurred at the end of a quarter or year).

With respect to the DOJ’s regressions, Oracle noted that the coefficient for SAP America was the same as the coefficient for PeopleSoft and concluded that the regression provided no support for the claim that PeopleSoft was a closer competitor to Oracle than was SAP America. Defendant’s use of the regression suggested that Oracle and PeopleSoft did not constitute a “localized node” in product space. Oracle also argued that the DOJ’s regression was improperly specified because it did not control for the number of competitors present in each competition.

Oracle ran an alternative regression that included dummy variables for PeopleSoft, JD Edwards, and SAP America and additional dummy variables for one competitor to Oracle in a given procurement (i.e., two competitors in total), two competitors to Oracle, and three competitors to Oracle (see Table 2-2). Oracle observed that SAP America and PeopleSoft continued to have very similar effects on Oracle’s discounts, that is, their regression coefficients were similar. The DOJ responded that Oracle’s regression was unreliable because the dummy variables for the number of competitors to Oracle were highly correlated with the dummy variables for the individual firms. Indeed, when the dummy variables for Siebel and niche competitors were included in Oracle’s regression, the dummy variables formed a linear combination, and the regression could not be estimated.

Oracle also criticized the plaintiffs’ use of customer declarations, claiming that they were subject to “selection” bias—the declarations were not from a random sample of Oracle’s customers but rather from a sample selected to help make the prosecution’s case. Oracle also claimed that by redesigning their products to be more attractive to the industries and customers that (hypothetically) would face increased prices after the merger,
firms such as SAP America and Lawson would defeat any price increase imposed by the merged firm. Finally, the defendant argued that it was not enough for the DOJ to establish that the proposed transaction was a 3-to-2 merger. In Oracle’s view, the plaintiffs had to establish that the merged firm would have dominance, given the differentiated nature of the relevant products, and the plaintiffs had failed to do so.

**Efficiencies**

The defendant noted that in hostile acquisitions the acquiring company cannot perform detailed analyses of synergies. Oracle argued, however, that the
merger would give rise to significant reductions in sales and marketing costs. Since these costs are variable, the savings would have a significant effect in disciplining prices after the merger.

**JUDGE WALKER’S DECISION**

Judge Walker ruled against the plaintiffs and in favor of defendant Oracle. His fundamental conclusion was as follows:

> The court finds that the plaintiffs have wholly failed to prove the fundamental aspect of a unilateral effects case—they have failed to show a "node" or an area of localized competition between Oracle and PeopleSoft. In other words, plaintiffs have failed to prove that there are a significant number of customers (the "node") who regard Oracle and PeopleSoft as their first and second choices. If plaintiffs had made such a showing, then the court could analyze the potential for exercise of monopoly power over this "node" by a post-merger Oracle or the ability of SAP or Lawson to reposition itself within the node in order to constrain such an exercise of monopoly power.

**Market Definition**

In his decision, Judge Walker noted that high-function enterprise software, as defined by the plaintiffs, had no recognized meaning in the industry. He concluded that there was no bright-line test to separate "large" customers from "mid-market customers," and he found that plaintiffs had not established that ERP vendors distinguished among these customers based on the amount they spent in an ERP purchase. Judge Walker summarized the characteristics of products sold by Oracle, PeopleSoft, and SAP America, Lawson, AMS, Microsoft, and best-of-breed vendors, as well as outsourcing services offered by firms such as Accenture and ADP, and found that the products were differentiated.

The judge noted that the plaintiffs’ case relied on customer witnesses, system integrator and industry witnesses, and their economic experts. The judge concluded that the testimony of the customer witnesses was largely unhelpful in defining markets for high-function HRM and FMS software because the customers offered little, if any, testimony on what they would do if Oracle increased prices postmerger. The judge concluded:

> There was little, if any, testimony by these [company] witnesses about what they would or could do or not do to avoid a price increase from a post-merger Oracle. To be sure, each testified, with a kind of rote, that they

*United States v. Oracle*, p. 1102.
would have no choice but to accept a ten percent price increase by a
merged Oracle/PeopleSoft. But none gave testimony about the cost of
alternatives to the hypothetical price increase a post-merge Oracle would
charge: e.g., how much outsourcing would actually cost, or how much it
would cost to adapt other vendors’ products to the same functionality that
the Oracle and PeopleSoft products afforded."

In contrast, the judge found the testimony of Oracle’s customer witnesses
more useful, since it was concrete and described specific actions that would
be taken if prices were raised.

The judge found that the DOJ’s market concentration statistics suf­
fered from several shortcomings. First, the DOJ based its calculations on
all transactions greater than $500,000. The judge concluded that the
sample of transactions considered was too small and that no attempt was
made to separate the HRM and FMS sales from the bundles in which they
were sold.

The judge noted that the DOJ stated that there was no "quantitative
metric" that could be used to identify high-function products, yet the DOJ
concluded that there was "something different" about the products at issue.
The judge concluded that plaintiffs had not proven that the relevant product
market was limited to high-function FMS and HRM software. He deter­
mined that products provided by outsourcing firms, mid-market vendors
such as Lawson and AMS, Microsoft (in conjunction with BearingPoint),
and best-of-breed solutions could not be excluded from the relevant prod­
uct market. The judge also found that legacy FMS and HRM systems (i.e.,
systems that a customer had in place prior to acquiring new systems) should
be excluded from the relevant product market. With respect to the geo­
graphic market, the judge found that the Elzinga-Hogarty test was a valid
method for establishing the geographic extent of the market and determined
that the test (as applied by the defendant) supported a worldwide market for
the products in question. Given his findings regarding the relevant product
and geographic markets, the judge concluded that the market share and
concentration statistics presented by the DOJ were inapplicable to the
antitrust analysis of the proposed transaction.

Competitive Effects

The judge observed that there is little case law on the analysis of the unilat­
eral effects of a merger. The judge cited the Horizontal Merger Guidelines
for a discussion of the "necessary elements of a unilateral effects claim
involving differentiated products under section 7." The Guidelines state

**United States v. Oracle, p. 1131.

* United States v. Oracle, p. 1117.
that two conditions must be shown for a finding of significant unilateral effects in a differentiated products merger:

Substantial unilateral price elevation in a market for differentiated products requires that there be a significant share of sales in the market accounted for by consumers who regard the products of the merging firms as their first and second choices, and that repositioning of the non-parties' product lines to replace the localized competition lost through the merger be unlikely. The price rise will be greater the closer substitutes are the products of the merging firms, i.e., the more the buyers of one product consider the other product to be their next choice.

The judge found this discussion incomplete because it emphasized the relative closeness of buyers' first and second choices and not the relative closeness of other alternatives in the market. He proposed four factors that together would constitute a differentiated products unilateral effects claim: (1) the products of the merging firms must be differentiated; (2) they must be close substitutes; (3) other products must be sufficiently different from the merging firms' products that a merger would make a small but significant and nontransitory increase in price (SSNIP) profitable; and (4) product repositioning by nonmerging firms must be unlikely. Of critical importance in the case, the judge elaborated on these conditions, stating as follows: "In a unilateral effects case, a plaintiff is attempting to prove that the merging parties could unilaterally increase prices. Accordingly, a plaintiff must demonstrate that the merging parties would enjoy a post-merger monopoly or dominant position, at least in a 'localized competition' space."

Against this conceptual background, the judge found that plaintiffs failed to establish the existence of a "node" or area of localized competition between Oracle and PeopleSoft. Specifically, plaintiffs had not established that a significant number of customers (the node) regarded Oracle and PeopleSoft as their first and second choices. Judge Walker criticized the plaintiffs for failing to use "thorough econometric techniques such as diversion ratios showing recapture effects" that might have met his four-prong test for unilateral effects.

The judge also commented on the three competitive effects analyses of the DOJ. In evaluating the plaintiff's analysis of Oracle's Discount Authorization Forms and regression analysis, which indicated that competition from PeopleSoft often resulted in larger Oracle price discounts, the judge noted that this evidence was offered in isolation and was not compared with similar information from PeopleSoft and SAP America Discount Authorization Forms. He concluded that the analysis established only that Oracle and PeopleSoft often competed vigorously with each

other. The analysis did not establish that this competition was any more intense than that between Oracle and SAP America or between PeopleSoft and SAP America. As such, Judge Walker concluded that the analysis did not establish the necessary node or localized competition space.

The judge concluded that the DOJ’s merger simulation model based on English auctions was unreliable because it was based on unreliable market shares, as he had previously determined in his analysis of the DOJ’s proposed market definition.

Finally, the judge noted that the plaintiffs presented no evidence at trial regarding coordinated effects, yet they included a section on coordinated effects in their post-trial brief. He noted that coordination would be difficult because the products were highly differentiated and because the market lacked price transparency. Given the absence of any evidence in the record regarding tacit market division, the judge concluded that the allegations of coordinated effects were without merit.

**Efficiencies**

The judge found that Oracle had not presented sufficient evidence to substantiate and verify its estimate of merger-specific efficiencies.

**CONCLUSIONS AND AFTERMATH**

There are two important points of disagreement between the DOJ’s position and the judge’s decision.

First, plaintiffs disagreed with the decision on the required elements of a unilateral effects case in a merger involving differentiated products. Plaintiffs argued, as set forth in the Horizontal Merger Guidelines, that two conditions must be met: “Substantial unilateral price elevation in a market for differentiated products requires that there be [1] a significant share of sales in the market accounted for by consumers who regard the products of the merging firms as their first and second choices, and that [2] repositioning of the non-parties’ product lines to replace the localized competition lost through the merger be unlikely.”

In contrast, the judge concluded that four elements were required to make such a showing: (1) the products of the merging firms must be differentiated; (2) they must be close substitutes; (3) other products must be sufficiently different from the merging firms’ products that a merger would make a SSNIP profitable; and (4) product repositioning by nonmerging

---

*United States v. Oracle*, p. 1169.

**“Horizontal Merger Guidelines**, at § 2.21.
firms must be unlikely. In addition, Judge Walker concluded that in order to
"prevail on a differentiated products unilateral effects claim, a plaintiff
must prove a relevant market in which the merging parties would have
essentially a monopoly or dominant position."

The judge’s characterization of the likely competitive effects of merg­
ers in differentiated products markets has been criticized by a number of
economists. They have argued that the judge’s conclusion that "a plaintiff
must prove a relevant market in which the merging parties would have
essentially a monopoly or dominant position" is incorrect in the context of
standard analyses of differentiated products mergers. As Werden (2006)
comments: "For a merger to produce significant price increases, the merg­
ing brands must be next-closest substitutes from the perspective of a sig­
nificant number of individual customers. But when viewed from the
perspective of all customers collectively, the merging brands need not be
especially close substitutes." Thus, the merging parties need not have
"essentially a monopoly or dominant position" in order to find it profitable
to impose a SSNIP in a differentiated products market. Shapiro (2005,
p. 15) comments that the judge "appears to be taking a Section 2 notion of
market power, namely monopoly power or dominance, and porting that
over to a Section 7 context. Economists recognize, however, that market
power is a matter of degree; there are weaker versions of market power than
monopoly power. [The judge] is conflating the two."

Moreover, if the judge’s third condition above were satisfied (i.e., if the
merging firms would profitably implement a SSNIP), then, by the smallest
market principle, the products of the merging firms would be a separate
antitrust market, and the combined firm would have a monopoly in the
"localized competition" space of interest, as well as a dominant position in
a relevant market. The additional conditions specified by the judge for
determining a postmerger monopoly or dominant position would then be
redundant. Under the conditions specified by Judge Walker, the combina­
tion would be a merger to monopoly, the prima facie case for banning the
merger would be met, and any additional analysis of unilateral effects
apparently would be unnecessary.

Second, the plaintiffs disagreed with the decision on how the presence
of price discrimination affects the analysis of unilateral effects. Plaintiffs
viewed the case as a "3-to-2" merger in which "competition in this case
involves a bidding process that is separate for each customer. . . . Because

"United States v. Oracle, p. 1123.
'Horizontal Merger Guidelines, at § 1.11. Market definition begins by defining relatively narrow
product and geographic markets and expanding those until a small but significant and nontransitory
price increase becomes profitable. The smallest market principle is that the relevant market is gen­
erally defined to be the smallest product and geographic markets for which the price increase
becomes profitable.
"Plaintiffs' Post-Trial Brief, p. 1.

Page 85
the price competition to sell high function HRM and FMS software is specific to the particular customer, the effects of the merger differ across customers according to the significance of the head-to-head competition between Oracle and PeopleSoft.” Thus, according to the plaintiffs’ position, each procurement was, from an economic perspective, an auction in which the price was set independently of other such auctions, in large part because arbitrage is impossible. The logical conclusion of this analysis is that each procurement contest or auction constitutes a separate market.

In contrast, the judge focused on the product differentiation aspect of the merger and criticized plaintiffs for failing to use “thorough econometric techniques such as diversion ratios showing recapture effects” that might have met his test for finding unilateral effects. However, in the context of independent auctions, the anticompetitive effects of a proposed merger could occur regardless of the diversion ratios between different products. The diversion ratio measures the fraction of sales lost by one product, when its price (to all customers) rises, that is subsequently recaptured by a second product, holding the price (to all customers) of the second product constant. In a price-discrimination market where “the price competition … is specific to the particular customer,” the merged firm can increase prices to individual customers by different amounts. In markets with such individualized pricing, the diversion ratio, which assumes that prices are increased to all buyers, plays no role.

An interesting aspect of the Oracle case is its implication for the role of market definition in future unilateral effects cases. Economic analyses of the competitive effects of mergers in unilateral effects cases can rely, as in the Oracle case, on simulations. As noted by Walker (2005) and Werden (2005), in principle merger simulations do not require that relevant antitrust markets be delineated. Indeed, in his opinion, Judge Walker stated: “Merger simulation models may allow more precise estimations of likely competitive effects and eliminate the need to, or lessen the impact of, the arbitrariness inherent in defining the relevant market.”

However, in practice, merger simulations often are calibrated to real markets, which can require (1) the identification of all relevant competitors and (2) their market shares, both of which require a delineation of the relevant market. As Budzinski and Christiansen (2007, pp. 155-156) conclude: “In practice, the ’inherently arbitrary’ task of delineating the relevant markets cannot automatically be avoided by the implementation of merger simulation models. Ironically, the court [in United States v. Oracle] rejected the simulation model exactly because [the court found] it did not consider all

---

"Plaintiffs’ Post-Trial Brief, p. 33.

*United States v. Oracle, p. 1122.

Page 86
relevant competitors and failed to include all the relevant products—in other words, the results of the simulation model were rejected because of a precedent inadequate market delineation." Thus, although in principle merger simulations reduce or eliminate the necessity to delineate relevant antitrust markets, the practical application of those methods appears to necessitate at least some consideration of relevant antitrust markets.

With respect to the role of customer testimony, the lesson of Oracle is that such testimony should be as detailed as possible on the subject of actions that customers likely would take in response to price increases. Customer testimony can be an effective complement to economic analysis. However, as Heyer (2007) observes, customers of merging firms may choose to be "rationally ignorant"—it may not be rational for customers to become knowledgeable about the ways in which the merger might harm them in the future, given the costs and benefits of acquiring that information. Also the economic incentives of direct purchasers, who may be manufacturers or distributors, may differ from those of final consumers. Moreover, customers may be reticent to state their views publicly since that could result in the disclosure of confidential business information.

REFERENCES


In May of 2003 Arch Coal, the third-largest coal producer in the Southern Powder River Basin (SPRB) area of Wyoming, announced that it would purchase Triton Coal Company, the fourth-largest producer of coal in the region. On April 1, 2004, the Federal Trade Commission (FTC) filed a complaint in federal district court seeking a preliminary injunction to block the transaction.

At the outset, the case itself was notable for the fact that the FTC’s sole theory of harm was a coordinated effects theory. The FTC argued that this merger would facilitate the exercise of “market discipline” by the three largest coal producers, who would jointly eliminate excess capacity that had persisted in the SPRB for well over a decade, which had kept spot prices near incremental costs and pushed down forward contract prices.

The decision by Judge John Bates reached conclusions in a number of areas that will no doubt lead to much discussion and debate among practitioners and antitrust scholars. The most notable is the finding that assertions by or concerns of customers that a merger will raise prices do not...

---

"The author was an FTC staff economist assigned to this case. The opinions expressed in this article are not necessarily those of the Commission nor any individual commissioner.

"The Merger Guidelines provide for both a unilateral effects theory and a coordinated effects theory. Under a unilateral effects theory, the merger gives the merged parties enough market power to raise prices above the competitive level on their own. Most recent cases have been brought under this theory. Under a coordinated effects theory, the merger is alleged to help facilitate tacit coordination among the remaining competitors.

See, e.g., Katz and Sheshinski (2007, p. 33); Froeb et al. (2005).
See, e.g., The Antitrust Source (2004).
constitute evidence that the merger will be anticompetitive and should be given little weight. A second included what some commenters have suggested is a much higher standard than had been used in the past for proving a coordinated effects case. This stems from an apparent unwillingness to accept a coordinated effects theory without proof that the parties had already successfully coordinated or the existence of a "theory for punitive discipline among producers" beyond a threat of expanding production in response to perceived deviation from tacit coordination by competitors. A third is that the decision stated that, "the case rests on a novel FTC theory of likely future 'tacit coordination' among competitors to restrict production, as opposed to direct coordination of prices." Finally, the decision concludes that total coal reserves owned by each company were a better measure on which to base the Herfindahl-Hirschman Index (HHI) calculation than was either actual production or production capacity.

BACKGROUND OF THE CASE

The SPRB is an area of northern Wyoming that contains a vast store of relatively low-BTU and very low-sulfur coal. At the time of the merger only five companies owned twelve surface mines that produced approximately 350 million tons of coal per year, which constituted about 40 percent of the steam coal produced in the United States. Depending on the mine, the coal, known as "sub bituminous coal," ranges from roughly between 8900 BTUs per pound to 8200 BTUs per pound. There is a thin secondary market on which two benchmark BTU levels are priced: 8800 and 8400. The five southernmost mines, known as "Tier 1" mines, produce 8800 BTU coal, and the other seven, known as "Tier 2" and "Tier 3" mines, produce 8400 BTU coal (though the Tier 2 mines have higher average BTU content than do the Tier 3 mines).

All of the mines in Tier 2 and Tier 3 produce coal that (when burned by electric utilities to produce steam for generating electricity) yields sulfur dioxide levels (SO₂) that average around .8 pound per million BTU, which is called "compliance coal" because it meets the 1990 Clean Air Act Phase II standards, without the use of scrubbers. Two of the 8800 mines also pro-

This conclusion is mirrored by the decision in U.S. v. Oracle Corporation, 331 F. Supp. 2d 1098 (N.D. Cal. 2004), brought at about the same time by the U.S. Department of Justice. See Case 3 in this book by McAfee, Sibley, and Williams.


Ibid., 115; emphasis added.

The Clean Air Act as Amended in 1990 requires that all electric utilities emit no more than 1.2 lbs of SO₂ per million BTU generated. Title IV of the Clean Air Act required under Phase I some of the most highly polluting utilities to emit no more than 2.5 pounds of SO₂ per million BTUs of heat generated by 1995. Phase II required a greater number of plants to emit no more than 1.2 pounds of
duce about .8 pound S\textsubscript{0.2}, and the other three mines produce levels closer to .4 lbs, which is known as “super compliant coal.”

The SPRB has a vast seam of coal that ranges between sixty and ninety feet thick. At the time of the trial this seam was between 150 and 250 feet below the surface, depending on the mine and surface topology. Thus, the strip ratio—the ratio of overburden (i.e., the ground that lies above the coal) to the thickness of the coal seam in the various mines—ranged from 2.5:1 to 4.1:1 at the time of the merger. The average cost of extraction in the SPRB increases by roughly $1 per ton for an increase of one point of strip ratio. The ratio increases slightly as mining operations move west over time. One estimate is that at 460 million tons per year mining rate, the weighted average strip ratio in the SPRB will increase from 2.9:1 in 2004 to just 3.2:1 in 2015. There are over twenty years of reserves remaining for which current mining technology is economical and a far larger number of years of reserves farther west at deeper depths.

Mining coal in the SPRB requires initially opening a large pit using large shovels called “drag lines” and a fleet of dump trucks and smaller shovels to expose the coal vein. Once exposed, a row of coal is dug by electric shovels and loaded into dump trucks, which dump the coal onto a conveyor system that carries the coal to a loadout, which is a large mechanism that is located above a railroad spur that loads coal onto trains that transport the coal across the country to the utilities that burn the coal to generate electricity. A mine’s output can be constrained either by truck/shovel fleet capacity (i.e., its capacity to dig coal) or its conveyor/loadout capacity (i.e., its ability to load coal onto trains). A mine typically has several truck/shovel fleets operating simultaneously.

Once a row of coal is removed, the trucks and shovels move west and begin removing the “next row” of overburden, placing it in the void left by the coal that was just mined. In this manner all SPRB mining operations are moving west. When a sufficient portion of a mine has been mined out, the land is reclaimed by returning the topsoil and returning the surface to a life-sustaining condition.

At the time of the merger it would have taken well over a year physically to open a new mine. This makes an existing pit a valuable asset

\$0. per million BTUs beginning in the year 2000. The amendment also established a “cap and trade” system, which allowed plants that were producing less than the allowed sulfur to sell sulfur credits to plants that were above the limit. Such a plan caps the overall level of sulfur emitted across the country but allows overall attainment costs to be lower than if each plant individually were required to meet the emissions requirements. Coals from the SPRB, the Uinta Basin, and Central Appalachia are the only coals that meet these standards. Other coals must be burned in facilities that have scrubbers installed in the smokestacks, which mix the exhaust with limestone and water to absorb the S\textsubscript{0.2}. Retrofitting utilities with scrubbers as well as operating scrubbers are very expensive.

\"As a basis for comparison, the mine-mouth price of coal was typically in the range of $4—$6 per ton.\"

\"Hill and Associates (2004, p. S-6). This would raise the variable cost of mining by about $0.30 per ton over this time period.\"
because it provides "easy" access to coal relative to opening a mine de novo.

The U. S. government owns virtually all of the land with coal reserves in the SPRB. The land is auctioned to the highest bidder under the "lease by application" (LBA) process. The winning bid is paid in five yearly installments independent of the rate of mining. At the time of the merger LBA bids ranged from about seventy-five cents per ton for the best coal to around twenty-five cents per ton for the least desirable coal. The LBA process for a given tract typically lasts around three to four years. Add to this approximately a year to reach coal physically, and de novo entry into this market (in the time frame considered relevant for antitrust analysis) is not a factor for antitrust analysis.

Purchasers of Coal

SPRB coal is used mostly by utilities in the Midwest, but utilities as far southeast of the SPRB as Florida and as far northeast as Maine purchase SPRB coal. Some have plants that were built specifically to burn the lower-BTU coal. Many other utilities were initially designed to burn higher-BTU, higher-sulfur coals from other regions of the country but switched to sulfur-compliant SPRB coal rather than incur the high cost of retrofitting scrubbers to meet the Clean Air Act requirements while burning higher-sulfur eastern or Illinois basin coal. The heterogeneity of utilities that burn SPRB coal causes it to be viewed as differentiated on at least four dimensions.

The first is utility location. For coal with a mine-mouth price between $4 and $6 per ton, transportation costs can represent up to 80 percent of the delivered costs. Since about 5 percent more 8400 coal must be purchased and shipped than 8800 coal to obtain a given total BTU level, utilities far from the SPRB value the 8800 BTU coal more highly than those closer to the SPRB.

The second is the BTU level's effect on boiler performance. Utilities designed to burn SPRB coal could operate at peak efficiency using either 8800 or 8400 BTU coal. However, for many utilities that had boilers that were designed to burn higher-BTU coals, the switch to SPRB coals reduced the output of their plants. In such cases 8400 BTU coal would mean lower power output than 8800 BTU coal.

12 Under this process, a mine owner (or anyone else) nominates a parcel of land. Once nominated, time is allotted for geological surveying and environmental studies, after which a sealed-bid auction is held, and the land is leased to the highest bidder, subject to the U.S. Bureau of Land Management's (BLM) reserve price—its estimate of the coal's "market value." Nominated parcels are typically located adjacent to existing coal mines and contain anywhere from five to fifteen years' worth of recoverable reserves for that mine.

13 Utilities purchase coal FOB the mine. They execute separate contracts for transporting the coal with the railroads.
The third is sulfur level. Different states have different rules regarding the levels of sulfur that could be emitted (beyond the Clean Air Act requirements). Further, states differed on the degree to which utilities could meet their sulfur requirements by buying sulfur credits and on whether utilities that were below their sulfur limits would be allowed to sell credits.

A fourth possible source of differentiation is that the Tier 3 mines were served only by the Burlington Northern Santa Fe (BNSF), whereas the Tier 1 and 2 mines were served by BNSF and the Union Pacific Railroad (UP) over the "joint line" that was co-owned by the two railroads. A utility with only a contract with the UP would have to pay an additional interchange fee to the BNSF to have coal delivered from Tier 3 mines.

**A Brief History of the SPRB**

Surface mining began in the SPRB in the early 1970s with the mining of 8400 BTU coal in what are now the Tier 3 mines. Because of the area’s very low strip ratios, the variable cost of production at this time was about $2 per ton (as reflected in the SPRB’s spot prices). As the mining moved west the strip ratios increased, and the cost of mining peaked at about $8 per ton in the early 1980s. At this point the technology improved (primarily in the form of larger trucks and shovels), and the cost of mining fell. By 1987, when the first 8800 BTU mines (Tier 1 mines) opened, the spot price of 8400 BTU coal had fallen to about $3.50 per ton, and the spot price of 8800 BTU coal was about $4.50 per ton. These prices persisted on average until 2000 (see Figure 3-1). During this time total SPRB production increased from 105 million tons to 350 million tons, suggesting no diseconomies of scale from expansion over this time.

During 1999 and 2000, a number of mines reduced production capacity. In January 2001, 8800 BTU coal spot prices spiked to $13 per ton (and 8400 BTU coal prices rose as well, peaking at about $3 less than 8800 BTU coal prices). In July of 2001 the Department of Justice announced an
TABLE 3-1
SPRB Mine Statistics

<table>
<thead>
<tr>
<th>Company</th>
<th>Mine</th>
<th>Tier</th>
<th>Avg. BTU 2003</th>
<th>SO² lb per mmBTU 2003</th>
<th>Strip Ratio 2004</th>
<th>Production 2003 mmtons</th>
<th>Loadout Capacity January 2004 mmtons/year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arch</td>
<td>Black Thunder</td>
<td>1</td>
<td>8819</td>
<td>.70</td>
<td>3.0</td>
<td>62.6</td>
<td>68 (idle)</td>
</tr>
<tr>
<td></td>
<td>Coal Creek</td>
<td>2</td>
<td>8328</td>
<td>.82</td>
<td>idle</td>
<td>idle</td>
<td>(idle (24))</td>
</tr>
<tr>
<td>Triton</td>
<td>N. Rochelle</td>
<td>1</td>
<td>8775</td>
<td>.48</td>
<td>4.0</td>
<td>23.9</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>Buckskin</td>
<td>3</td>
<td>8457</td>
<td>.81</td>
<td>2.0</td>
<td>17.5</td>
<td>24</td>
</tr>
<tr>
<td>Peabody</td>
<td>N.A.R.C.</td>
<td>1</td>
<td>8838</td>
<td>.45</td>
<td>3.2</td>
<td>80.1</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>Caballo</td>
<td>2</td>
<td>8511</td>
<td>.82</td>
<td>2.4</td>
<td>22.7</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Rawhide</td>
<td>3</td>
<td>8236</td>
<td>.64</td>
<td>1.0</td>
<td>3.6</td>
<td>10</td>
</tr>
<tr>
<td>Kenncott</td>
<td>Antelope</td>
<td>1</td>
<td>8839</td>
<td>.57</td>
<td>2.5</td>
<td>29.5</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Jacobs Ranch</td>
<td>1</td>
<td>8719</td>
<td>1.02</td>
<td>3.0</td>
<td>35.5</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>Cordero Rojo</td>
<td>2</td>
<td>8475</td>
<td>.78</td>
<td>3.0</td>
<td>36.1</td>
<td>48</td>
</tr>
<tr>
<td>R.A.G.</td>
<td>Belle Ayr</td>
<td>2</td>
<td>8583</td>
<td>.68</td>
<td>3.8</td>
<td>17.8</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>Eagle Butte</td>
<td>3</td>
<td>8413</td>
<td>.83</td>
<td>2.1</td>
<td>24.7</td>
<td>30</td>
</tr>
<tr>
<td>Totals for Active Mines</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>354</td>
<td>421</td>
</tr>
</tbody>
</table>


investigation of the major producers in this region for price fixing. No case was ever filed. By 2002, 8800 BTU coal spot prices had fallen to $5.50, and in 2003 spot prices had risen to about $6.50.

In 1984 each of the three tiers produced about 35 million tons. By 2003 Tier 1 production had risen to 232 million tons, Tier 2 production had risen to about 76 million tons, and Tier 3 production had risen to just 46 million tons, for a total of 354 million tons.

Consolidation from 1993 to 2000 reduced the number of SPRB owners from ten to five, who at the time of the merger consisted of Arch Coal, Triton Coal Company, Peabody, Kenncott, and R.A.G. (now Foundation Coal). The relevant information concerning these companies and their mines can be found in Table 3-1.

Market Definition

There was no serious claim by either side that the market included any coal produced outside the SPRB. This is because two of the four other major coal-producing regions—northern Appalachia and the Illinois Basin—produce coal that does not meet the Clean Air Act sulfur requirements, so no plant that used SPRB coal but did not have scrubbers could switch to coals.
from these regions. Although the coals from the other two regions—central Appalachia and the Uinta Basin (Utah)—were sulfur compliant, they were much more expensive to mine and therefore had prices that were too high to allow them to be economical substitutes for SPRB coal if there were a 10 percent increase in SPRB prices for most utilities using SPRB coal.

Thus, the only question was whether 8800 BTU coal was sufficiently differentiated from 8400 BTU coal to constitute a separate market. The merger guidelines prescribe the SSNIP test, in which a set of products is assumed to be controlled by a monopolist, and one asks, if that monopolist were to raise the price of those products by 5 to 10 percent, would enough customers continue to purchase to make the price increase profitable? If such an increase is profitable, then the relevant market is no larger than that set of products under consideration. The smallest such set of products is usually deemed to be the relevant antitrust market.

The transactions data in this market come primarily from reports that utilities file with FERC, in which they report coal as it is delivered. These data do not allow for a standard demand analysis that could be used to calculate the cross-price elasticity between 8800 and 8400 BTU coal because utilities contract for most quantities (about 80 percent) more than one year in advance, and even “spot” quantities can be contracted for up to a year in advance of the delivery date. Thus, these data cannot be used to determine relative prices of different coals at the time that each purchase decision is made and therefore cannot be used to calculate cross-price elasticities.

Another data source in principle could have been used to calculate the cross-price elasticity between 8800 and 8400 BTU coal. The defense produced a witness who explained how a utility evaluates bids by computing the effective cost per kilowatt generated for each coal. This analysis effectively computes a utility’s cross-price elasticities among the competing coals. Amassing such analysis from the fewer than two hundred utilities that purchased SPRB coal at the time of the merger would have provided a good basis for performing the SSNIP test.

Neither side presented results based on such analysis, and therefore neither could present direct evidence on the market’s willingness to substitute between 8800 and 8400 coal. Instead they presented more indirect evidence. The FTC’s economics expert provided some co-integration analysis and argued that there was sufficiently low correlation between the spot price of 8800 and 8400 BTU coals to suggest that these were in different markets but that he could not conclude this definitively from this evidence. **The FTC also**

---

**Co-integration analysis looks at the correlation between the prices of two products over time. If the prices are highly positively correlated, then they are likely to be substitutes; and if there is little correlation between the prices, then it is likely they are not close substitutes.**

**Rather than analyze market definition in-house, the FTC subcontracted this analysis to an outside consulting firm that was able only to estimate incentives to switch coal types based on transportation costs but was unable to account for differences in sulfur levels or effects on boiler productivity of lower BTU coal. Consequently this analysis had little value and was not presented at trial.**
TABLE 3-2
SPRB Coal Market Concentration (HHI)

<table>
<thead>
<tr>
<th></th>
<th>Practical Capacity</th>
<th>Loadout Capacity</th>
<th>Production</th>
<th>Reserves</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current Market</td>
<td>2152</td>
<td>2068</td>
<td>2201</td>
<td>2054</td>
</tr>
<tr>
<td>Postmerger</td>
<td>2346</td>
<td>2292</td>
<td>2365</td>
<td>2103</td>
</tr>
<tr>
<td>Postmerger Increase</td>
<td>193</td>
<td>224</td>
<td>163</td>
<td>49</td>
</tr>
</tbody>
</table>


Presented the testimony of Triton’s director of marketing, who stated that he did not consider the price of 8800 BTU coal when bidding on RFPs for Triton’s 8400 BTU coal from the Buckskin mine, nor did he consider the price of 8400 BTU coal when setting prices for North Rochelle’s 8800 BTU coal.

Defendants’ expert presented an analysis that used estimated differences only in transportation costs across utilities to generate a cross-price elasticity between 8800 and 8400 BTU coal. She found that a 10 percent price increase of 8800 BTU coal only would not be profitable. However, on cross-examination she had to admit that a table she presented showed that a 25 percent price increase of 8800 BTU coal would be profitable.

Seller Concentration
The FTC provided calculations based on measures of productive capacity such as loadout capacity and practical capacity, as well as on actual production. Defendants argued that the best basis was total reserves of coal that each mine owner had under lease in the SPRB. As is indicated in Table 3-2, all of the measures yielded an HHI level of a little over 2000. The FTC’s measures yielded increases in HHI of 163-224, while the defendants’ measure yielded an increase of just 49.

THE FTC’S CASE
The FTC brought its complaint based entirely on a “coordinated effects” theory, eschewing any notion of “unilateral effects.” It proposed that the

“Under a coordinated effects theory the merger is alleged to be anticompetitive because it will increase the chances that some set of firms can successfully engage in tacit coordination. A unilateral effects theory alleges that the merger will allow the newly merged firm to raise prices on its own.
three largest owners—Arch, Peabody, and Kennecott (dubbed the "Big Three")—would have a greater incentive to restrict output and raise price once Triton’s 8800 BTU North Rochelle mine, which was positioned to expand output in the face of decreased output on the part of the Big Three, was acquired.

First, the FTC argued that the merger would increase the profitability of a coordinated price increase by the Big Three and so made coordination—even without any explicit (illegal) communication—more likely. North Rochelle produced some of the SPRB’s highest-value coal because of its high BTU and very low sulfur content. It had the largest unused load-out capacity in the SPRB, making future expansion inexpensive. It had been the fastest expanding mine in the region over the previous five years and had publicly stated that because of its excess loadout capacity, it was positioned to expand output rapidly.

Thus, the theory was that if the Big Three had tried to restrict its output to raise prices, Triton could expand North Rochelle production (since the mine was cash-flow positive”), significantly reducing the Big Three’s profitability. This in turn made it difficult for each of the Big Three to expect significantly higher prices if they reduced output, making tacit coordination difficult. Once Triton was acquired, its threat of expanding would be eliminated, and the FTC’s economic expert provided analysis showing that the two remaining fringe firms would be less likely to expand output (which would drive prices down) and that a tacit agreement among the Big Three to compete less aggressively, postmerger, would be much more profitable.

Second, the FTC maintained that the market characteristics facilitated tacit coordination. Output is observable, since the U.S. Department of Labor’s Mine Health and Safety Administration (MHSA) publicly reported production at each mine on a quarterly basis along with the number of worker-hours used to produce that coal. Also each regulated utility is required to report each shipment quantity, delivered price, BTU level, sulfur level, and ash level to the Federal Energy Regulatory Commission (FERC) and the U.S. Department of Energy’s Energy Information Administration (EIA), which made this information public. Since each owner knew the mine-mouth price it charged its own customers, it could back out the rail rates that each of its customers paid. It could then subtract these rates from the delivered prices its customers reported from competing SPRB mines to obtain the mine-mouth prices charged by competitors. Thus, it would be difficult to cheat on a tacit agreement by secretly altering the quantities, prices, or characteristics of the coal because those characteristics would be reported

The FTC argued that it was mainly proceeds from North Rochelle that were being used to pay down Triton’s sizeable debt.

within three months or less of delivery. Mines also had access to COALdat\textsuperscript{20} data, which provided estimates of mine-mouth prices for spot sales and longer-term contract sales.

The market was characterized by many small transactions, and quantities and delivered prices were publicly available. The mines were located in the same area and were open-pit mines. Any increase in production, productive capacity (such as additions of truck and shovel fleet), or loadout capacity would easily be noticed. Thus, firms could observe the production of their competitors, which would facilitate detection of cheating.\textsuperscript{21}

Third, the owners of the Big Three mines had publicly stated their need for all producers to exercise discipline and reduce output. In response to the persistent low spot prices, Arch's CEO Steven Leer had on several occasions, along with the CEOs of Peabody and Kennecott, made public statements at industry trade shows that low prices for coal were due to the mines' having too much capacity and their willingness to sell "incremental tons" at a price just above cost and that therefore each mine should reduce its capacity. In 1999 and 2000 each of the Big Three publicly announced at trade shows that they had reduced production capacity and encouraged other companies to do the same. In 2001 during the price spike CEO Leer stated publicly that the price increase was due to the reduction in capacity made in the previous years and that ownership consolidation facilitated this discipline. In 2002 Leer publicly announced that Arch was reducing its output even though this would reduce Arch's short-run profits because he believed that the positive profits he would make on these sales were not high enough. The FTC argued that such public statements were a means by which the Big Three could coordinate output reductions.

In April of 2003 Leer privately affirmed that the strategy of withholding uncommitted tons from the market was the correct strategy but that he did not know how long Arch could continue to "lead this charge." He warned that if prices did not improve soon, Arch would ramp up the mines to full production.\textsuperscript{22} The FTC argued that such expansion was a

\textsuperscript{20}COALdat is produced by Platts, which provides information on energy prices (including the data upon which Figures 3-2 and 3-3 are based).


\textsuperscript{22}On direct examination Leer offered an explanation that differed from the FTC's assertion that "lead this charge" meant leading a tacitly coordinated output reduction in the SPRB. He stated that "ramping up production" applied to a general company wide strategy, and not just the SPRB. With respect to the phrase "lead this charge," Leer testified, "It's really referring to a discussion I had with the board at a prior board meeting when we were making a projection of 2003 and thinking that 2003's market was going to be improving and that we were going to try to go after long-term contracts as opposed to short-term spot sales. The board agreed with that strategy, and in that presentation and discussion—and it was really a discussion of the board—I made the point that we were all getting on this horse together, and if the strategy didn't work off [sic], we would all fall off the horse together. And when in fact we get to April of 2003, the market had not improved. We had mis-read the market. And I'm telling the board that we have fallen off our charger . . . ." Trial Transcript 6-28-04 am, p. 41.
mechanism by which deviation from a coordinated output restriction could be punished.

Fourth, the FTC argued that the Big Three had already begun to coordinate successfully. Specifically, the Big Three had engaged in “production discipline” and “price discipline.” It pointed to the fact that in late 2000 spot prices for 8800 BTU coal increased from $4.50 per ton to a high of $13 per ton and then fell over the next nine months to roughly $5.50 per ton, then increased to about $6.50 per ton and fluctuated in this range until the time that the merger was challenged.

The public statements by the Big Three along with the price spike led to complaints by utilities that immediately preceded the Justice Department’s announced investigation of price fixing in July 2001.

Fifth, Arch’s planned spinoff of Triton’s Buckskin mine to a non-SPRB mining company (in response to the FTC’s concerns) would not constrain tacit coordination because Buckskin produced the third-lowest BTU coal in the SPRB with some of the highest sulfur levels, which was not a good substitute for Tier 1 mines’ coal. Buckskin’s 2003 output was essentially the same as its 1998 output despite the increase in the SPRB output of over 20 percent during that time period. The FTC argued that Buckskin was not a profitable mine that could constrain the Big Three. Arch purchased the two Triton mines for a total of $364 million and then sold Buckskin for only about $80 million.

Finally, the FTC argued that a vast majority of customers (by coal volume) believed that the merger would cause prices to increase.

For its affirmative case the FTC called as witnesses a number of purchasers of coal for electric utilities. They explained why they could not or had very limited ability to switch from SPRB coal to other coals outside the region. They all testified that they had heard the public statements by the CEOs of the Big Three and were worried that this would lead to higher prices. They explained that even though they had some long-term contracts, these contracts almost always had clauses that required prices to adjust to market prices every three to four years, so increases in current prices affected both the prices of coal that they were going to buy and coal that they had already committed to purchase. They also testified that there had been consolidation in mine ownership in the past that had not caused concern. This merger, however, made them believe prices would increase, in part because of the statements made by the Big Three owners that they should all produce less coal, and in part because utilities typically purchased coal from several different sources, and this merger reduced the number of sources of 8800 BTU coal from four to three.


The FTC also called a representative from RAG whose testimony included the statement that while RAG had plans to expand output, it would not do so if it meant reducing the price of coal.

The FTC’s economics expert explained many of the key points of the FTC’s case listed above, including how the merger increased the Big Three’s profitability from coordination and that the SPRB market had many characteristics that made it susceptible to coordinated interaction.

THE DEFENDANTS’ CASE

The defendants made the following arguments: First, the merger would combine two adjacent mines, which would create an efficiency that would lower the cost of production.

Second, this case was unprecedented since this was a "5-to-5" merger. Triton operated two mines: North Rochelle and a Tier 3 mine, Buckskin. During the investigation Arch contracted to sell Buckskin to Peter Keiwit and Sons, a mining company that had no presence in the SPRB, if the merger were to be consummated. Thus, the number of mine owners would not fall.

Third, the FTC was presenting an untested legal theory. Defendants stated that the FTC was arguing that the Big Three were coordinating on quantity. They argued that this was materially different from past coordination cases, which were based on price.

Fourth, defendants argued that it would be difficult for the Big Three to coordinate. Such coordination, they argued, would be difficult because mines had the ability secretly to sign long-term contracts for coal, undercutting any supracompetitive prices, and because demand was increasing, making it very difficult for the Big Three to agree on the quantity levels on which to coordinate.

Fifth, even if the Big Three did attempt to restrict output to raise price, the three remaining mines—Eagle Butte and Belle Ayr mines owned by R.A.G. and the to-be-spun-off Buckskin, all 8400 BTU mines (dubbed the "fringe firms")—would expand output sufficiently to prevent prices from rising to supracompetitive levels. They argued that 8400 and 8800 BTU coal were sufficiently interchangeable so that 8400 BTU coal could constrain 8800 BTU coal prices.

Sixth, rather than being a maverick, Triton since May of 2001 had adopted an "antimaverick" strategy. Defendants argued that Triton adopted a "last mine standing" strategy in which it responded to requests for proposals (RFPs) from utilities by bidding prices in excess of the competitive level and waiting for the other SPRB mines to run out of capacity, at which point, Triton could make sales at supracompetitive prices.

Seventh, defendants argued that the reductions in production capacity that the FTC argued constituted coordinated interaction were actually
unilateral decisions made in response to firms' realizing that they had excess capacity.

Finally, the defense argued that the market was not concentrated enough for the merger to be harmful. In particular the defense argued that HHI calculations should be based on the total reserves of SPRB coal owned by the mines at the time of the merger rather than on a measure of output, or productive capacity, and that under a measure based on reserves the increase in the HHI was relatively small—only 49—and thus did not warrant antitrust intervention.

The defense’s case focused on several themes. One was that no customer had any knowledge that the mines were attempting to coordinate. Utility officials who took the stand were asked if they had any specific knowledge of mines attempting to coordinate, to which the answer was "no." Many were asked if they had any knowledge of specific mines’ cost structures, which again received negative answers. Many were also asked if they believed the market was competitive, which was answered affirmatively.

A second theme was to assert that (with respect to market definition) for most utilities 8800 and 8400 BTU coals were good substitutes and that the choice between them was simply a matter of price; thus the relevant market was the broader market of the two coals combined. One important defense witness was a coal purchaser for a utility who produced a bid evaluation sheet in which the utility took each bid for a specific request for proposal (RFP) and converted it into an effective cost per kilowatt hour generated. Thus, the differences in the coals were reduced to the differences in the cost of producing electricity.

A third theme was that historical data were not a good basis on which to predict what would be bid today. An important defense witness was the director of marketing for one of the parties who testified that he would not use three-month-old price data as the basis to choose what to bid on current spot market contracts. Similarly, witnesses were asked if current spot prices were a good indicator of future long-term prices.

A fourth theme was that Triton was too "weak" to constrain the Big Three’s pricing and that Buckskin was Triton’s strong mine, while North Rochelle was the weak mine. The key witness here was the CFO from Triton, who testified that the poor financial condition of Triton had resulted in debt covenants, which prevented it from offering competitive prices. Instead, North Rochelle was forced to adopt a "last mine standing" strategy.

**THE DECISION**

The decision began by defining the market and identifying the change in the HHI as a result of the merger. It then addressed two major questions. The first was the degree to which coordination could occur in the SPRB.
The second was which mines not owned by the Big Three owners would prevent the Big Three mines from successfully coordinating.

**Market Definition**

The decision found that the relevant market was all SPRB coal. The court observed that the FTC had the burden of proving an 8800 BTU coal market. It focused only on the FTC’s expert, noting that he was not willing to state that the evidence definitively implied that 8800 BTU coal was a separate market. The decision turned to the only thing left, which was testimony by several utilities (including some of plaintiffs) that the choice between 8800 and 8400 BTU coal was a matter of economics. “Based on the reluctance of the plaintiff’s own expert to conclude that 8800 BTU SPRB coal is a separate relevant market and the significant interchangeability between 8800 BTU and 8400 BTU coal, the court declines to find that 8800 BTU coal is a separate relevant market.”

**HHI Analysis**

The Court’s decision regarding market concentration focuses on the “best” basis for calculating HHI. The decision first references *General Dynamics,* a merger case in which the acquired coal mine had already sold all of the reserves it owned and had no hope of acquiring any more reserves. In that case, the Supreme Court held that because the acquired firm was selling no new coal, it had no impact on competition. The Arch Coal decision quoted *General Dynamics,* “... A more significant indicator of a company’s power effectively to compete with other companies lies in the state of a company’s uncommitted reserves of recoverable coal.”

The decision then cites the FTC’s expert as saying that the North Rochelle mine had 7.5 years of reserves left and no impediment to purchasing additional reserves through the LBA process and states that, “Notwithstanding the Supreme Court’s reliance on reserves in *General Dynamics,* exclusive reliance on that measure to determine market concentration in the SPRB today does not appear warranted. The present SPRB coal market is different from the relevant coal market in *General Dynamics*—particularly in the use of more short-term contracts. Moreover the parties have presented statistics based on all reserves, but only uncommitted reserves were considered in *General Dynamics* and are particularly relevant to assessing future competitiveness. The greater availability of additional reserves through LBAs has not been fully accounted for in the parties’ presentations based on present reserves. Hence although the Court


will look primarily to market concentration as measured by reserves, it will also consider other measures."

Cooperation in the SPRB

The "Novel Theory"

The decision describes the FTC’s theory of competitive harm as a "novel theory" based on coordination only on output in which "the major coal producers in the SPRB market would constrain their production so that increases in supply would lag behind increases in demand, thereby creating upward pressure on price . . . What this means is that the FTC must show projected future tacit coordination, which itself may not be illegal, which is speculative and difficult to prove, and for which there are few if any precedents." The decision goes on to explain, "As defendants have noted, prior coordinated effects challenges to mergers based on alleged output coordination have invariably been accompanied by a coordinated effects theory grounded on price coordination." Further it stated, "the novel approach taken by the FTC in this case makes its burden . . . more difficult."

The FTC’s subsequent appeal of the Court’s decision highlighted the facts that output reductions and price increases are two sides of the same coin and that there was nothing "novel." The appeals court agreed. However, the court did not find this to be sufficient to overturn the district court’s ruling.

Likelihood of Coordination

The first major question was whether the Big Three would be more likely to coordinate after the merger. The decision cites the merger guidelines by noting that "successful coordination requires two factors: (1) reaching terms of coordination that are profitable to the firms involved and (2) an ability to detect and punish deviations that would undermine the coordinated interaction. . . . Credible punishment . . . may not need to be any more complex than temporary abandonment of the terms of coordination by other firms in the market."

In describing what the FTC would need to show, the decision states "... plaintiffs must establish that the proposed transactions will increase the risk of coordinated output restriction and decrease the likelihood of deviation.

---

28 Ibid., 131.
29 Ibid., 131. Emphasis in original.
30 Ibid., 132.
in the SPRB market. This requires a sophisticated attempt to show a developing propensity towards this form of tacit coordination in the SPRB market supported by an ability of the SPRB firms to monitor each other’s behavior.” In describing what the court would look for, it states, “The Court will proceed by examining the competitive state of the SPRB market today, determining whether coordinated interaction is feasible and, if so, whether there is evidence that actual or tacit coordination has occurred, and then examining the structure and dynamics of the SPRB market, the competitive strength of Triton, and the likely roles that RAG and Kiewit would play in a post-merger market.”

Here again, it may be that this decision “raises the bar” significantly as to what must be shown when bringing a coordinated effects case. Some might interpret “increase the risk of coordinated output restriction” as showing that the profitability of coordination is increased by the merger,” but there is almost no discussion of incentives in the decision despite the FTC’s expert testimony that the merger would raise the profits from coordination.

Instead the judge seems to have focused on whether coordination had already taken place.” While the decision does not state that a showing of past coordination was a requirement for this court to find harm, the judge seems to have placed great weight on this fact. If other courts adopt this posture, that would raise the bar. Such a standard would clearly permit mergers that could create coordination where previously there had been competition, which is untenable. Thus, showing past coordination can be a “plus factor” for showing the likely harm that would arise from a merger, but it cannot reasonably be a prerequisite for bringing a coordinated effects case.

Analysis of Past Coordination

The decision concluded that there was no history of successful coordination in the SPRB. The FTC had argued that in 1999 and 2000 each of the Big Three had idled productive capacity and in public statements at trade shows had urged its competitors to do the same. Then in 2001 spot prices for 8800 BTU coal increased from $4.50 per ton to a peak of $13 per ton and then fell back to about $6 per ton.

The decision found that the idling of productive capacity was not evidence of past coordination because each incident of capacity reduction could be explained as a unilateral decision to save costs in a time of overcapacity. (Though the decision does suggest that statements at trade shows about these

*Ibid., 132.
*See Kovacic et al. (2006).
*The FTC may well have pushed the court in this direction by stating that it would show that the firms had already successfully coordinated.
unilateral decisions were "indicative of possible producer coordination.""

The price spike in 2001 was the result of unusually low stockpiles at utilities and of a number of unforeseen events such as severe weather, an increase in the price of natural gas, and several eastern mines closing due to bankruptcies. Finally, quoting the defendants' economics expert, the high prices were explained by, "... Buyers reacted to these various factors by resorting to panic buying."

The decision also noted that a number of customers, when asked on the stand, testified that the market for coal was competitive.

Susceptibility of the Market to Coordination

The decision finds that "although production restrictions were advocated and even practiced by Arch during 2000-2002, and broader coordination by SPRB producers to limit supply was feasible, no express or tacit coordination to limit production has actually occurred among the major SPRB coal producers."

The decision next argues that the structure and dynamics of the SPRB market do not make coordination likely. The reasons were the following:

The products in the SPRB are heterogeneous and "heterogeneity of products and producers limit or impede the ability of firms to reach terms of coordination.""

"FERC data are reported on a delivered basis ... and do not break out the cost of coal alone."

"FERC statistics (and MSHA data) also do not address a number of critical items and provide no information on most of the competitively significant terms and conditions of coal contracts ... For example FERC data do

\[ FER v. Arch Coal, Inc., 329 F.Supp. 2d 109, 137. \]

\[ Ibid., 133. \]

\[ See The Antitrust Source (2004, p. 10) for a discussion of whether customers have the expertise to conclude whether the market was competitive or not. \]

\[ FTC v. Arch Coal, Inc., 329 F.Supp. 2d 109, 140. \]

\[ The decision cites a facilitating practices case, In re High Fructose Corn Syrup, 295 F.3d 651 (7th Cir. 2002) (HFCS), in which high fructose corn syrup manufacturers had in place a mechanism that facilitated maintaining supra-competitive prices. The decision states that many of the factors present in HFCS were not present in the Arch case; \]


\[ Expecting factors in a facilitating practices case to be present in a coordinated effects merger case seems to raise the bar unrealistically, since a successful facilitating practices case alleges that there is a successful mechanism that keeps prices high (without a merger), while a coordinated effects case posits an imperfect mechanism that becomes more effective due to the merger. \]

\[ Ibid., 140. \]

\[ Ibid., 141. \]

The decision apparently believed that it was impossible to back out the mine-mouth price from the delivered price using the analysis described by the FTC's expert economist.
not capture the various terms and conditions of coal contracts that affect the overall value of the contract to the producer.""

FERC data are incomplete because 20% of utilities are not required to file. The pricing data are not timely: "... it is simply not possible to estimate current coal prices by looking at FERC or (MSHA) data."

“Demand for SPRB coal is not predictable either in the short- or the long-term. The two largest demand drivers for coal consumption are the weather and the economy, and neither can accurately be predicted”

[Citing defendants’ expert] "The opaqueness of utility inventories also makes it difficult to predict the demand for SPRB coal from year to year."

"The emphasis on sealed bids and confidentiality is an important aspect of the market structure and dynamic that would frustrate coordination among producers."

This aspect of the decision is summarized in the following passage:

Tacit agreement would also be difficult to coordinate in this market place because terms of agreement would be hard to communicate between producers... Moreover there is no effective mechanism in the SPRB to discipline any producer that would deviate from the terms of coordination. Plaintiff’s economic expert posited no theory for punitive discipline among producers. Due to the nature of the confidential bidding and contracting procedures that gives producers incentives to submit aggressive bids to capture long term contracts, cheating would not be detected until well after the fact, if ever, and any punishment would come well after the fact as well.”

This passage may imply an increase in the burden to show that coordination is likely. It seems to reject the idea that reduced prices due to a breakdown in tacit coordination once cheating is detected can deter defection from coordinated behavior, even though this specific mechanism is outlined in the Merger Guidelines. Rather, the decision seems to be searching for an instant retaliation response.

"Ibid., 141. These passages do not indicate what competitively sensitive terms other than price, quantity, BTU level, sulfur level, and delivery date are not captured in these data.

"Ibid., 142. Here the decision apparently takes the leap that because one does not use historical prices to estimate spot prices, one cannot use these data to infer whether competitors have been offering low prices in the past.

"Ibid., 142.

"Ibid., 142.

"Ibid., 144.

"Ibid., 145.
The Fringe Mines Could Constrain the Big Three

The decision states that, "It is defendants' burden to show that the 'fringe' of the SPRB market would expand sufficiently to defeat any merger induced price increase." The decision finds that Kiewit and R.A.G. were both "strong" companies with "credible plans to expand production and lessen the risk of coordination . . . The record evidence establishes that they could and likely would expand production to meet any production constraint by the major SPRB producers."

The decision credits the three fringe firms with being able to expand sales by an additional 31 million tons if 8800 BTU coal prices increased, even though they jointly produced just 60 million tons in 2003; and, while they accounted for 20 percent of SPRB output in 1998, they accounted for just 2 percent of the 65 million-ton increase in SPRB output between 1998 and 2003.

North Rochelle as Maverick

In deciding whether North Rochelle (the mine that Arch would retain) could act as a constraint on the Big Three, the court had to decide whether to consider North Rochelle's behavior during the period before 2002, in which it expanded output, and was identified by Arch as an impediment to production discipline, or to consider its behavior after 2002, the period after the time "they agreed to have Arch try to acquire them" when it ceased to be competitive and adopted its "last mine standing" strategy. The court chose to focus primarily on North Rochelle's behavior after 2002.

Customer Concerns

Perhaps the portion of this decision that will have the greatest effect on future cases is its treatment of customer concerns over the merger. The FTC produced several customers who expressed the opinion that the shrinking number of suppliers would lead to higher coal prices (as well as affidavits from a vast majority of customers with the same opinion). The decision found that "...the substance of the concern articulated by the customers is little more than a truism of economics: a decrease in the number of suppliers may lead to a decrease in the level of competition in the market. Customers do not, of course, have the expertise to state what will happen in the SPRB market, and none have attempted to do so. The Court therefore concludes that the concern of some customers in the SPRB market that

"Ibid., 147-148. However, there is no discussion of any evidence indicating that the only two fringe firms would expand enough to keep prices at competitive levels.

"Trial Transcript 6-24-04 am, p. 39.
transaction will lessen competition is not a persuasive indication that coordina
tion among SPRB producers is more likely to occur."

It is reasonable to assert that customer conclusions about the struc
ture of the economic theory of harm is likely to be beyond a customer’s expe	rise and that economic expert testimony is needed to explain how cer	ain data in the market should be interpreted in light of the theories of harm under consideration. However, the general statement, “Customers do not have the expertise to state what will happen in the SPRB market,” raises the question of what customers can say about the future of prices following a merger. This question is highlighted by this exchange between the judge and the FTC’s counsel during opening statements:

COURT: How much weight should I give to a customer who is saying this is what I think will happen?

FTC: . . . [T]hese individuals are involved directly in the purchase of coal. Their job is to project what the company, their own company, will be paying in the future to scout the industry, see what prices are, see what the state of competition is, can we get a better deal from another coal supplier . . .

Such considerations are highlighted by the testimony of one customer who testified that in response to the merger announcement he (on behalf of the utility he represented) offered to purchase coal from Arch under a ten-year contract at a price that was significantly above the rate of return that Arch was currently earning and offered to lock in this rate of return by tying price increases to changes in Arch’s cost. The negotiations did not result in a contract as Arch wanted prices tied to changes in the market price instead of its own costs. This testimony raises the question of how much weight customer conclusions about future price increases should be given in light of the facts that the conclusion was reached in the ordinary course of business and that the utility was willing to commit resources (offering prices in excess of current competitive levels) based on this conclusion.

At the end of the day the judge had heard that the majority of cus
omers believed that this merger would lead to increased prices and that sellers said increased concentration had already led to increased prices. Yet, the court found after its own examination of the testimony that this increase in concentration would not lead to an increase in price.

“ibid., 146.

“Trial Transcript 06-21-04 am, p. 35.
THE MARKET SINCE THE DECISION

The decision was made public on August 13, 2004. The FTC immediately appealed, requesting an emergency stay of the decision. The appellate court responded in August with the following two-sentence decision: "Ordered that the motion be denied. Although the court agrees that there is nothing novel about the theory it [the FTC] has advanced in this case, the court concludes it has not met the standard for an injunction pending an appeal."

The merger was immediately consummated. In June 2005 the Commission announced that it would not pursue the case any longer.

The performance of the market after the merger is quite interesting. It is impossible to reach a conclusion regarding the effects of the merger or the efficacy of tacit coordination without a detailed analysis of both spot prices and longer-term contract prices, as well as any changes in the costs of production. That would require access to nonpublic information beyond the scope of this chapter. However, the behavior of the market since the decision raises some interesting questions regarding the level of competition in the SPRB.

Figure 3-2 shows the spot prices for each of the major coal regions, including SPRB 8800 BTU coal, between July 2004 and July 2007. Between August of 2004 and December 2004 this spot price hovered between $5.75 and $6.50 per ton. Then between December 2004 and May 6, 2005, the price increased from $5.75 per ton to $8.01 per ton.

Analysis of the spot prices after May 2005 is complicated by the fact that on May 14 and 15 of 2005 severe rainstorms washed away several sections of the joint line track serving the SPRB. This and the subsequent program of track maintenance over the next several months led to a significant reduction in the ability of utilities to send trains into the SPRB to receive delivery of coal. Between May and November 2005 the Union Pacific (UP) railroad invoked a Force Majure clause and delivered essentially only about 85 percent of the coal that it had contracted to deliver.

(Interestingly, the BNSF invoked similar clauses for only a few weeks after the washouts.)

As Figure 3-2 indicates, in the months following the washout, the spot price of 8800 BTU coal increased to a peak of about $17 per ton in October and then to about $20 in January of 2006. After this the price fell steadily to just under $10 in September of 2006. Between September of 2006 and

57 See, e.g., http://www.spreservices.info/PDF/06/06-055_u_57_l.pdf at 3.
July of 2007, the price fluctuated between $8.50 and $10, remaining far above the premerger spot price, which was in the range of $5.75-$6.50 per ton.\footnote{See \url{http://www.eia.doe.gov/cneaf/coal/page/coalmkt/coalmar.html}.}

This pricing behavior, especially just after the rail washout, creates some interesting analysis issues. The public information available indicates that the trains that are allowed into the SPRB are controlled by the railroads and the utilities. The mines have no control over which trains (and therefore which customers) pull coal from the SPRB. Thus, during the period in which the railroads were the bottleneck, the demand for spot coal from the perspective of the SPRB mines would have fallen. That is, there would be fewer trains and therefore fewer tons to be purchased relative to the expected demand for the mines served by the joint line (which included the Tier 1 and Tier 2 mines).\footnote{If the quantity of coal that each utility could extract from the SPRB was constrained by its allocation of train capacity, and there was more capacity in the production of coal than train capacity, then the spot market between...}
mines and utilities were, say, perfectly competitive, and coal were essentially homogeneous, one would expect the spot price to fall, as the mines in the SPRB, experiencing excess production capacity, compete the price down.”

The rail outage created a scarce resource (which is rail capacity). The resulting scarcity rents should accrue to those who control the scarce resource, which in this case is either the utilities or the railroads (depending on whether existing contracts allowed the railroads to raise rates in response to the shortage). However, as Figure 3-2 shows, the price of 8800 BTU coal increased (and though not shown in the figure, the price of 8400 BTU coal increased as well, but the difference between the two prices increased to about $2.50 per ton).

Once the bottleneck problems were eliminated, one would expect that there would be an increase in demand for SPRB coal. During the bottleneck period the utilities would have drawn down their inventories, and once the bottleneck was eliminated they would be expected to demand more coal than they burned in order to build up their inventories. Thus, once the bottleneck was eliminated one could expect higher prices in the SPRB as demand would exceed the mines’ short-run capacity at spot prices near average variable cost of production. This would seem consistent with the $20 spot price in January 2006 and the reduction in the price over the next several months as utilities restocked their inventory levels.

A more detailed analysis might observe that the rail shortage, in reducing the tonnage that could be removed from the SPRB, also reduced the total amount of BTUs that could be extracted from joint line mines. In an attempt to increase the amount of BTUs, utilities at the margin would substitute away from the lower-BTU coal in the joint line mines (i.e., the 8400 BTU coal) toward the higher-BTU coal mines (i.e., the 8800 BTU mines). However, this does not appear to have happened. In the three quarters in 2005 during which the rail shortage occurred, the Mine Health and Safety Administration website reports that production from the four 8800 BTU mines actually fell by 5 million tons relative to the same three quarters in 2004. At the same time, production from the 8400 BTU mines in those three quarters of 2005 increased by about 3 million tons.

Thus, it is difficult to reconcile the spot market price increases as being caused by the rail shortage and the SPRB behaving as if it were perfectly competitive to gain by bidding up the price of coal, since increasing what it was willing to pay could not increase the quantity of coal it could purchase. Thus, it is difficult under such a fact scenario to ascribe spot price increases to the increased willingness to pay of utilities.

To the extent that there are secondary market transactions, one would expect the secondary market prices to be substantially above the competitive price in the primary market, as there is no excess productive capacity in the secondary market. Effectively the supply of coal for physical delivery on the secondary market would be constrained by any rail capacity constraint.

(or at least "very") competitive. Again one needs to be careful regarding this analysis because there are institutional details that are not available, such as the composition of coal that was already committed under long-term contracts during the last three quarters of 2005 and contractual limitations on the ability of utilities to switch at the margin. The decision anticipates large output increases from the three fringe mines in response to increased prices from a merger-induced supply reduction. ("Hence fringe expansion by Kiewit and RAG would be enough to cover demand [sic] shortfall... and defeat any price increase.") One might also expect an increase in production from the two Tier 3 fringe mines in 2005, which are not served by the joint line, and from all three fringe firms in response to the high 8800 BTU spot prices in 2006 during a period in which utilities are presumably trying to replenish inventories. Since the merger was consummated in September of 2004, virtually all of the coal produced in 2004 would have been committed before the merger. Between 2004 and 2006, Buckskin increased its output by 2.5 million tons to 22.8, Eagle Butte increased its output by 2.4 million tons (though

TABLE 3-3
SPRB Production (millions of tons)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;Fringe Mines&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Buckskin</td>
<td>17.3</td>
<td>17.5</td>
<td>20.3</td>
<td>19.6</td>
<td>22.8</td>
</tr>
<tr>
<td>Eagle Butte</td>
<td>18.1</td>
<td>24.7</td>
<td>23.0</td>
<td>24.1</td>
<td>25.4</td>
</tr>
<tr>
<td>Belle Ayr</td>
<td>22.5</td>
<td>17.8</td>
<td>18.7</td>
<td>19.3</td>
<td>24.6</td>
</tr>
<tr>
<td>Other 8400 Mines</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rawhide</td>
<td>5.3</td>
<td>3.6</td>
<td>6.9</td>
<td>12.4</td>
<td>17.0</td>
</tr>
<tr>
<td>Coal Creek</td>
<td>7.1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3.1</td>
</tr>
<tr>
<td>Caballo</td>
<td>26.0</td>
<td>22.7</td>
<td>26.5</td>
<td>30.5</td>
<td>32.8</td>
</tr>
<tr>
<td>Cordero Rojo</td>
<td>37.0</td>
<td>36.1</td>
<td>38.7</td>
<td>37.8</td>
<td>39.7</td>
</tr>
<tr>
<td>Tier 1 Mines</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black Thunder</td>
<td>42.7</td>
<td>62.6</td>
<td>72.2</td>
<td>87.6</td>
<td>92.6</td>
</tr>
<tr>
<td>N. Rochelle*</td>
<td>0.4</td>
<td>23.9</td>
<td>15.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N. Antelope/Rochelle*</td>
<td>64.6</td>
<td>80.1</td>
<td>82.5</td>
<td>82.7</td>
<td>88.5</td>
</tr>
<tr>
<td>Antelope</td>
<td>19.4</td>
<td>29.5</td>
<td>29.7</td>
<td>30.0</td>
<td>33.9</td>
</tr>
<tr>
<td>Jacobs Ranch</td>
<td>29.1</td>
<td>35.5</td>
<td>38.5</td>
<td>37.3</td>
<td>40.0</td>
</tr>
<tr>
<td>Total</td>
<td>289.7</td>
<td>354.0</td>
<td>372.2</td>
<td>381.3</td>
<td>420.4</td>
</tr>
</tbody>
</table>

between 2003 and 2004 Eagle Butte output fell by 1.7 million tons) to 25.4, and Belle Ayr’s output increased by 5.9 million tons to 24.6, for a total production of 72.8 million tons (see Table 3-3).

In November of 2004 Hill and Associates published its annual report on the SPRB. In this report Hill and Associates predicts the output of each of the mines. For 2006 Hill and Associates predicted output levels of 24.25, and 22 million tons for Buckskin, Eagle Butte, and Belle Ayr, respectively, for a total of 71 million tons.” Recognizing that is just a prediction of one consultant,” one might nonetheless consider it an unbiased estimate of expected production levels that did not account for the rail problems in 2005. With a predicted output of 71 million tons and an actual output of 72.8 million tons, virtually all of the output of the three fringe mines was anticipated in advance of the rail outage in June of 2005. Therefore, it would seem difficult to conclude that the three fringe firms generated much of a response at the margin to the price increases in 2005 and 2006, as one might have expected based on the decision.

CONCLUSION

The Arch Coal case is an interesting case both because it is one of very few recent cases that has relied exclusively on a coordinated effects theory of harm and because of a number of very strong conclusions by the district court about the level of competitiveness of the market and the likelihood of coordination. These conclusions will no doubt be debated by antitrust scholars. And as time passes and more data and analyses of the SPRB become available, it will be interesting to see how well they match the findings of the decision.

REFERENCES


"See Hill and Associates (2004, Table S-1).
"Defendants' and plaintiff's experts quoted from Hill and Associates reports at trial.
"Between 2004 and 2005 Rawhide increased production from 6.8 million tons to 12.4 million tons. Hill and Associates (2004) predicted that Rawhide would increase output to 12 million tons in 2005. It also predicted that in 2006 Rawhide would increase its production capacity to 17 million tons. In 2006 Rawhide did produce 17 million tons.


CASE 4

Sky Wars: The Attempted Merger of EchoStar and DirecTV (2002)

Richard J. Gilbert and James Ratliff*

INTRODUCTION

On October 28, 2001, EchoStar Communications Corporation announced its intention to acquire the assets of Hughes Electronics Corporation. At the time of the proposed merger, EchoStar operated the Dish Network, and DirecTV was operated by Hughes, which was then a wholly owned subsidiary of General Motors. EchoStar and DirecTV were the only direct broadcast satellite (DBS) companies that provided multichannel video programming distribution (MVPD) services to the entire continental United States.

Satellite television service in the United States dates from the late 1970s. The first satellite television services operated at relatively low power in the C-band of the electromagnetic spectrum at roughly 4 GHz and required a receiving antenna roughly four to eight feet in diameter. "DBS service" generally refers to higher power transmissions in the Ku-band at about 12.2 to 12.7 GHz. The higher power and shorter wavelength of these transmissions (wavelength is inversely related to frequency) made possible the use of much smaller receiving antennae than those used in C-band systems.


*The authors consulted for the Department of Justice in its investigation of the proposed EchoStar-DirecTV merger.
was a medium-power technology and required a receiving dish about three feet in diameter. DirecTV launched its higher-power all-digital DBS service in June 1994. DirecTV’s service required a receiving dish the size of a large pizza, smaller than PrimeStar’s dish and far smaller than the trailer-sized dish required for C-band reception. PrimeStar was not competitive with the DirecTV service, which offered more programming with a smaller dish antenna. DirecTV purchased PrimeStar in 1999 and migrated all PrimeStar subscribers to DirecTV equipment. EchoStar’s Dish Network launched its DBS service in the United States in March 1996, two years after DirecTV. The receiving dish formats are similar for EchoStar and DirecTV, although the two systems are not compatible because they use different signal encryption methods.

Sales of DBS systems mushroomed and rapidly displaced sales of C-band systems. DirecTV and EchoStar each had more than two million subscribers within two years from their launch dates. DirecTV, with its two-year head start and acquisition of the PrimeStar installed base, was the larger of the two, but EchoStar closed the gap with a combination of aggressive pricing and promotion. By 2002, DirecTV had grown to more than 10.9 million subscribers, and EchoStar had more than 7.5 million subscribers (DOJ 2002a).

DBS networks broadcast from satellites placed in geosynchronous orbits, so that the satellites appear in a fixed position in the sky when viewed from the Earth. International Telecommunications Union rules determine the number and location of these orbits to avoid interference between transmissions. There are three orbital locations that provide DBS coverage over the entire continental United States excluding Alaska (called "full-CONUS slots") located at 101, 110, and 119 degrees west longitude (W.L.). The FCC governs assignment of the orbital positions and available transponder frequencies to DBS carriers. Thirty-two transponder frequencies are available at each orbital slot, totaling ninety-six full-CONUS frequencies. Each frequency can carry multiple channels (e.g., ESPN and HBO).

Several companies controlled the ninety-six available full-CONUS frequencies in the early years of the DBS industry, but by mid-1999 DirecTV and EchoStar had acquired rights to all ninety-six frequencies, with EchoStar controlling fifty and DirecTV controlling the remaining forty-six. Although high-power DBS service can also be offered from non-full-CONUS slots, these slots are less desirable for a mass-market DBS service because they do not cover the entire continental United States. At the time of the proposed merger, EchoStar had capacity for 500 channels, and DirecTV had capacity for 460 channels in standard definition under the assumption of 10 channels per transponder.

*A transponder receives the uplink signal from a ground station and rebroadcasts the signal to the ground.*
Both the Antitrust Division of the Department of Justice (DOJ) and the Federal Communications Commission (FCC) reviewed the proposed merger of EchoStar and DirecTV. The FCC was involved because the merger would have required the transfer of various broadcast licenses and authorizations for which the FCC has authority. One year after EchoStar and DirecTV announced their intent to merge, the DOJ, joined by the attorneys general of twenty-three states, the District of Columbia, and the Commonwealth of Puerto Rico, filed a civil antitrust lawsuit to block the proposed acquisition. The DOJ said that if the merger were allowed to proceed, it would eliminate competition between the nation’s two most significant DBS services and substantially reduce competition in the MVPD business to the detriment of consumers throughout the United States. A few days earlier the FCC had announced its objection to the proposed merger and ordered the matter set for an administrative hearing. EchoStar and DirecTV abandoned their proposed merger in response to the challenge by the DOJ and concerns indicated by the FCC.

While mergers typically challenge antitrust enforcers to balance possible efficiency benefits against the risk of higher prices, the proposed merger of EchoStar and DirecTV was extraordinary in both respects. As the only two DBS suppliers in the United States, the combination of EchoStar and DirecTV was a potential merger to monopoly for consumers without access to cable television. For most other consumers, the merger would lower the number of suppliers of multichannel video programming distribution services from three to two, although the merging parties argued that competition from cable would be sufficient to discipline pricing by the merged satellite supplier. The proposed merger also had the potential to create significant efficiencies. There was a large overlap in the programming supplied by EchoStar and DirecTV, and both companies faced capacity constraints. A merger of the two DBS suppliers could eliminate much of this duplication and free up capacity for other services. The dilemma was whether the merger-specific efficiencies outweighed the possible harm from higher prices.

The analysis by the DOJ and the FCC followed the methodology described in the DOJ and FTC Horizontal Merger Guidelines (DOJ and FTC 1997). We first discuss product and geographic market definition. Next we turn to competitive effects, which, along with efficiencies, were a central issue for analysis of the proposed merger. We also discuss the potential for entry and claimed efficiencies, and we close with a retrospective look at the performance of EchoStar and DirecTV since their proposed merger was abandoned.
PRODUCT MARKET DEFINITION

EchoStar and DirecTV are in the business of multichannel video programming distribution (MVPD). According to the FCC, MVPD suppliers are entities that offer multiple channels of video programming for purchase by subscribers or customers. The services that the FCC considers to be within MVPD include cable, direct broadcast satellite (DBS), multichannel, multipoint distribution service (MMDS), satellite master antenna television (SMATV), and C-band.

MVPD subscribers represented 86.4 percent of all TV households in 2001 (FCC 2002b, Table C-1). Table 4-1 shows each service's share of MVPD subscribers in 2001. MMDS, SMATV, and C-band together accounted for less than 4 percent of all MVPD services in 2001, and their share was not increasing. The DBS share has grown substantially and accounted for 75 percent of the growth in MVPD households from 1997 to 2001. Cable systems can be digital or analog. Digital cable offers higher quality, more channels, and other content such as greater choice of pay-per-view movies. DirecTV and EchoStar are all-digital services.

Product market definition was not a particularly contentious issue in this proposed merger. While the parties and the antitrust enforcement agencies at times disputed the precise boundaries of the relevant product market, these disagreements had little impact on the general conclusion that a merger of EchoStar and DirecTV would substantially increase concentration in markets that were already highly concentrated.

The merging parties asserted that the relevant product market included all MVPD services and may be even broader (Willig 2001, p. 4), although the parties' competitive analysis focused on DBS and cable. The DOJ complaint alleged that the relevant product market affected by the proposed merger was multichannel video programming distribution but rejected C-band service as an acceptable substitute due to the high cost and inconvenience of the necessary equipment (DOJ 2002a). Given the small share of MVPD services other than cable and DBS, whether these other MVPD services were included in the relevant market had little quantitative significance for the effects of the proposed merger on estimated market shares. The DOJ also noted that digital cable systems are closer substitutes for DBS than are older analog cable systems.

The FCC adopted a MVPD product market for the purpose of its Hearing Designation Order. However, the FCC also found evidence to

---

37 C.F.R. § 76.1000(e).

*MMDS is a broadband wireless technology. SMATV is generally intended for distribution of TV signals to households in one or more adjacent buildings. Other technologies, such as DSL, were not available at the time of the proposed merger to provide MVPD services and were not widely deployed several years later.*
support a narrower market definition. The FCC noted that: MVPD services are highly differentiated; the evidence strongly suggests that EchoStar and DirecTV are closer substitutes for each other than for services of cable systems or other MVPDs; and high-capacity digital cable systems appear to be a closer substitute for DBS than are low-capacity analog cable systems (FCC 2002a, pp. 49-51).

Most consumers do not consider over-the-air broadcast television to be an acceptable substitute for cable or DBS services, particularly the many consumers who live in areas with poor over-the-air reception. Over-the-air broadcast television does not include the variety of programming that is available to cable and DBS subscribers. It does not provide the number of channels or access to popular channels such as ESPN or CNN and premium services such as HBO or Showtime. Despite the fact that over-the-air programming is available for free, most consumers willingly pay hundreds of dollars per year to subscribe to cable or DBS service. As was noted above, as of June 2001, more than 86 percent of all U.S. TV households subscribed to a fee-based MVPD service. Furthermore, cable prices have been increasing over time, while over-the-air broadcasts have remained free, which suggests that competition from over-the-air broadcasts is not sufficient to discipline cable prices. Applying a Merger Guidelines test, it is highly unlikely that a small but significant price increase by a hypothetical cable/DBS monopolist would be made unprofitable by customers switching to over-the-air television, which is reason to conclude that broadcast television is not in the same relevant market as cable and DBS.

A credible case could be made that DBS and digital cable are sufficiently distinct from analog cable to be in a separate relevant product market. Both DBS and digital cable have higher quality video and audio compared with analog cable and over-the-air broadcasts and offer more programming content. DBS, as an all-digital service, is more closely comparable with digital cable than with analog cable. Digital cable commands a premium price relative to analog cable. At about the time of the proposed merger, most cable systems required that consumers subscribe to a digital tier on top of their regular analog service in order to receive digital broadcasts.

### TABLE 4-1 National MVPD Subscriber Shares (June 2001)

<table>
<thead>
<tr>
<th>Percent of MVPD Subscribers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cable                     78.0</td>
</tr>
<tr>
<td>DBS                       18.3</td>
</tr>
<tr>
<td>SMATV                     1.7</td>
</tr>
<tr>
<td>C-band                    1.1</td>
</tr>
<tr>
<td>MMDTS                     0.8</td>
</tr>
</tbody>
</table>

Source: FCC Tenth Annual Report (2004), Table B-1.
In July 2001, the average monthly price of the most highly subscribed digital tier for cable systems was $9.62, compared with $33.81 for the most highly subscribed expanded tier that included only analog channels. This implies an average monthly price for digital cable service equal to $43.43, which is a premium for digital service of about 28 percent. The significant price gap between the price of analog and digital cable service indicated that there was considerable differentiation between the two services. Separate market-share calculations were performed corresponding to a product market composed of DBS and all cable systems and a market composed of DBS and only digital cable systems.

Cable and DBS services are not uniformly available to all consumers. Estimates of the fraction of households passed by cable in 2001 ranged from about 81 to 97 percent, but many homes that were passed by cable did not have access to digital cable. According to the FCC, about 76 percent of all cable subscribers had access to digital cable in 2001 and 88 percent in 2002 (FCC 2003b, Table 10). DBS is not an option for some consumers. DBS requires an unobstructed line of sight between the satellite and the customer’s dish. Some consumers either do not have an unobstructed line of sight to the satellite or do not want to place a dish in a location with an unobstructed line of sight, and others live in dwelling units that do not allow the placement of DBS equipment.

GEOGRAPHIC MARKETS

The merging parties initially contended that the relevant geographic market is national in scope because both EchoStar and DirecTV have national pricing plans (Willig 2001, p. 11), although the parties’ competitive analysis subsequently accounted for local differences in competition. While the DBS companies provide nationwide services, cable is provided locally. Even if EchoStar and DirecTV programming services were priced uniformly for all customers in the continental United States, the effects of a merger of the two nationwide DBS suppliers would depend on cable prices and service offerings at the local level. Furthermore, both EchoStar and DirecTV have targeted promotions at local levels and have the ability to price at a local level if they choose to do so. Structural analysis of the merger was performed at the level of the cable franchise area for regions that have cable service and at the Designated Market Area (DMA) level for regions without cable service.

---

5 FCC (2003b). These prices exclude equipment costs, which were somewhat higher for digital service.
7 A Designated Market Area is a geographic area defined by Nielsen Media Research as a group of counties that makes up a particular television market.

Page 120
Nearly all of the MVPD markets were highly concentrated, and the Herfindahl-Hirschman Indices (HHIs) for the proposed merger exceeded the structural screen in the DOJ/FTC Horizontal Merger Guidelines for identifying mergers that potentially raise significant competitive concerns. FCC staff computed concentration indices for geographic markets corresponding to 4984 local cable systems. The median postmerger HHI was 5653, and the median increase in the HHI was 861 for local MVPD markets consisting of DBS and all cable systems, whether analog or digital. When the product market was limited to DBS and high-capacity (digital) cable systems, the median postmerger HHI was 6693, and the median increase in the HHI was 206 (FCC 2002a, p. 57). These figures likely underestimate the significance of the proposed EchoStar-DirecTV merger for the structure of MVPD markets because DBS was experiencing rapid growth in the time period leading up to the proposed merger. Additional growth would increase the shares of EchoStar and DirecTV and increase the change in concentration from the merger.

The market definition analysis suggests concern that a merged EchoStar and DirecTV would have sufficient market power to raise prices above premerger levels. This concern is apparent whether the relevant product market includes all cable systems or is limited to the subset of digital cable systems. In regions—primarily rural areas—that do not have access to cable, the merger of EchoStar and DirecTV would change the structure of the market from two MVPD competitors to a single monopoly provider of MVPD services. As noted above, the estimated percentage of households without access to cable ranged from 3 percent to 19 percent in 2001, depending on the source of the data. An even higher percentage of DBS subscribers are in regions that do not have cable access. Thus, for a large number of DBS households the merger of EchoStar and DirecTV would be a merger to monopoly. Many other households do not have access to digital cable, and the merger of EchoStar and DirecTV would create a monopoly supplier of digital MVPD services for these households.

COMPETITIVE EFFECTS

While market shares identify structural conditions that could allow the merged firm—"New EchoStar"—to exercise market power, they do not demonstrate that the merged firm would raise prices or increase quality-adjusted prices by reducing programming offerings or other dimensions of service quality.

"This ignores the very small population receiving C-band and MVPD services other than cable and DBS in rural areas.

'The effect would not be a merger to monopoly for all of these households because some might not be able to receive DBS signals, and some might not have had access to both EchoStar and DirecTV before the merger.
For households with cable access, it is possible that New EchoStar's incentives to compete with cable would be sufficient to restrain postmerger prices. Furthermore, if New EchoStar continued to price DBS service on a national basis, then competition from cable where cable is a competitor might provide sufficient discipline to prevent price increases for households that are without cable access.

Proponents of the merger argued that the merger would not adversely affect price or service quality because EchoStar and DirecTV compete to attract customers from cable more than from each other. However, evidence of direct competition between EchoStar and DirecTV. Historically, EchoStar and DirecTV have offered similar prices and services. Figure 4-1 shows the monthly price of popular programming packages offered by EchoStar and DirecTV. EchoStar charged a lower price over this period, although the difference never exceeded $2 per month.

Expanding the number and diversity of channels is another way to lower the effective price. Figure 4-2 shows the price per channel for EchoStar and DirecTV for their most popular programming packages. See, e.g., Willig (2001).

"Total Choice was DirecTV's most popular package during this period. AT50, AT60, and AT100 were EchoStar's most popular packages at different times during this period. At every point in..."
The price per channel decreased from $0.55 in early 1998 to about $0.35 to $0.38 by the beginning of 2000 and remained at that level for the next two years.

Competition between EchoStar and DirecTV appears to have been particularly intense for customer premises equipment and installation prices, which dropped from several hundred dollars to near zero. The DOJ complaint noted that EchoStar itself acknowledged in papers filed in U.S. district court in 2000 (in an antitrust suit that EchoStar had brought against DirecTV, which was eventually dropped) that DirecTV and EchoStar reacted primarily to each other when setting equipment and service prices. The DOJ complaint also offered evidence from EchoStar CEO Charles Ergen regarding the importance of head-to-head competition with DirecTV. In an e-mail concerning Alaska, where DirecTV does not have a strong signal, Ergen wrote: "[r]ealize we have signal in Alaska and DirecTV doesn’t have much, we don’t have competition there. We don’t need to be as aggressive."

Another indication of competition between EchoStar and DirecTV was the total price of subscribing to these services, which appeared to be slightly lower than the price of cable after adjusting for digital versus analog service. Comparison of DBS and cable prices is complicated by the time, the graph reflects the price per channel of the EchoStar package that was most popular at that time.

These figures exclude installation and equipment fees.

"MacAvoy (2002, pp. 34-35) describes several instances of competitive responses by each DBS provider to price and quality offerings by the other company."
specific composition of programming packages, differences in equipment and installation costs, and the cost and availability of other services. In July 2001, the average monthly rate for cable was $33.81, excluding equipment costs (FCC 2003b, Table 1). This bought an average of about fifty-nine analog channels, including local broadcast networks. Digital cable service, when it was available, required an additional fee. In 2002, EchoStar offered about sixty all-digital channels for a monthly fee of $31.99. Further complicating the comparison is that local channels for EchoStar cost an additional $5.99 per month, and equipment costs could differ for multiple television sets. As previously noted, DirecTV prices were close to the prices charged by EchoStar.

Even if DBS prices were unambiguously lower than cable prices, this does not allow a definitive conclusion that competition between EchoStar and DirecTV is more important than competition between DBS and cable because lower DBS prices could be driven by competition between DBS and cable for MVPD customers. Certainly, a large fraction of the potential customers for DBS were people who were actual or potential cable subscribers. Whether New EchoStar would have continued to have an incentive to price low to win market share from cable requires a more thorough analysis of the demand for DBS and cable services.

Coordinated and Unilateral Effects

A merger can result in higher prices by facilitating coordinated interactions in the markets affected by the merger or by increasing the ability of the merged firm to profit through unilateral increases in price. Coordinated effects arise when firms in the industry recognize their mutual interdependence and take actions that sustain their joint profits. The DOJ complaint alleged that the merger presented a risk of coordinated effects because the merged company could target discounts and promotions to regions of the country in which the local cable company departed from a collusive pricing scheme.

Unilateral effects arise when the merged firm has sufficient market power to raise prices above premerger levels with no accommodating price response from other firms in the relevant markets. While the competitive analysis of the proposed EchoStar-DirecTV merger recognized the risk of coordinated effects, much of the analysis by the parties, the DOJ, and the FCC focused on the likely unilateral effects from the merger.

The ability to increase prices unilaterally depends on the elasticity of demand for the products sold by the merged firm and the extent to which the products of the merging firms are differentiated from each other. Under the assumption that the merged firm maximizes profit, the elasticity of demand for DBS and the marginal cost of supplying DBS services
determine the postmerger price of DBS through the necessary condition for profit-maximization:

\[
\frac{P_j - MC_j}{P_j} = -\frac{1}{\epsilon_j^{\alpha}}
\]  

(1)

In equation (1), \( P \) is the price of product \( j \), \( MC \) is its marginal production cost, and \( \epsilon \) is its own-elasticity of demand. The left-hand side of (1) is the Lerner Index of market power.

Analysis of the demand for DBS is complicated by the fact that there is little variation in the price of DBS systems. EchoStar and DirecTV charge a single national price for each service package, and that price has moved little over time. There is some variation in effective prices from changes in programming and from price promotions, but the changes in programming were also nationwide, and it is difficult to obtain accurate and comprehensive data on price promotions.

Goolsbee and Petrin (2004) measured DBS price elasticity by exploiting symmetry conditions in demand. Specifically, if a household’s share of expenditures on all MVPD program is independent of prices, then the effect on household demand of a $1 increase in the price of DBS is equivalent to the effect of a $1 decrease in the prices of all substitute MVPD services. Using this relation, they estimated the demand elasticities in Table 4-2.\(^{14}\)

The services in the Goolsbee and Petrin (2004) demand analysis are over-the-air broadcasts (antenna), expanded basic cable, premium cable, and DBS. Expanded basic cable is an analog service with about sixty channels including local networks. Premium cable service includes one or more premium networks such as HBO. The analysis aggregates demand for the two DBS providers, EchoStar and DirecTV.

\(^{14}\)The estimated price elasticities for DBS in Table 4-2 are lower in magnitude than the estimates in a previous working paper by the same authors that was available at the time of the merger (Goolsbee and Petrin, 2001). Independent analysis performed by the authors of this chapter and others using confidential data resulted in estimated price elasticities for DBS that were closer to those in Table 4-2.

<table>
<thead>
<tr>
<th>Service</th>
<th>Expanded Basic</th>
<th>Premium Cable</th>
<th>DBS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antenna</td>
<td>1.30</td>
<td>.92</td>
<td>.12</td>
</tr>
<tr>
<td>Expanded Basic</td>
<td>-1.54</td>
<td>.92</td>
<td>29</td>
</tr>
<tr>
<td>Premium Cable</td>
<td>1.26</td>
<td>-3.38</td>
<td>49</td>
</tr>
<tr>
<td>DBS</td>
<td>.93</td>
<td>1.17</td>
<td>-2.45</td>
</tr>
</tbody>
</table>


\[\text{TABLE 4-2} \]

Estimated Price Elasticities (Goolsbee and Petrin, 2004)
The elasticities in Table 4-2 have the expected signs. Own-elasticities are negative, and cross-elasticities are all positive, implying that the different types of services are substitutes, although only weakly so for DBS and over-the-air broadcasts.

The estimated elasticity of demand for DBS implies that a merged DBS company would have the ability to raise price significantly above marginal cost. Using equation (1) and assuming the own-elasticity $\epsilon_{DBS} = -2.45$ for DBS service estimated by Goolsbee and Petrin (2004), a merged DBS company would maximize its profit by choosing a Lerner Index of about 40 percent. This corresponds to setting price about 70 percent above marginal cost.

The demand analysis showed that the profit-maximizing price for New EchoStar would be significantly above its marginal cost, but the analysis did not establish that the merger would lessen competition and result in higher prices, adjusted for the quality of service, relative to premerger prices. Given an accurate estimate of marginal cost, the DBS elasticity is sufficient to forecast the postmerger profit-maximizing price, which can be compared with premerger prices to estimate the likely price increase from the merger. However, marginal cost is difficult to measure, in part because some cost components vary only for some customers. For example, installation and equipment costs are sunk for existing customers, but they are marginal costs for the acquisition of new customers.

Using accounting data on programming expenses and customer acquisition costs, MacAvoy (2002, p. 41) estimates monthly marginal costs of $26.80 for DirecTV and $30.39 for EchoStar. From equation (1), the -2.45 own-elasticity of demand for DBS in Table 4-2 implies a postmerger price of $44.20, using the lower marginal cost estimate for DirecTV, or $50.12, using the higher marginal cost estimate for EchoStar. Both estimates are significantly higher than the premerger prices for EchoStar and DirecTV. These prices were roughly $30.99 per month for EchoStar and $31.99 per month for DirecTV, plus $5.99 per month for local channels, excluding fees for premium services. Equipment and installation were typically free, although customers could be charged extra for special installation requirements or additional set-top boxes.

An alternative approach employed in modern unilateral effects analysis is to estimate the matrix of own- and cross-price elasticities of demand for products in the relevant markets and use these estimates, along with assumed competitive behavior, to estimate marginal production costs. This approach generates estimates of the premerger Lerner Indices for EchoStar and DirecTV. Given actual prices, the premerger Lerner Indices can be used to estimate marginal production costs. The estimated marginal costs then permit estimates of likely postmerger prices using equation (1).

In theory, the matrix of own- and cross-price elasticities can be estimated using a refinement of the econometric approach followed by
Goolsbee and Petrin (2004) in which consumers first choose between DBS and other MVPD services and then choose between EchoStar and DirecTV. In this “nested” choice model, the nest parameter measures the similarity of consumer tastes for the two DBS services and therefore the competition that would be lost if EchoStar and DirecTV were to merge. Unfortunately, experiments with this approach were inconclusive. With very little variation in the prices charged by EchoStar and DirecTV, it was difficult to obtain robust estimation results for the nest parameter. Instead, we consider other, more heuristic approaches to estimate the loss of competition from the merger.

Two approaches are useful to estimate whether the merger would likely result in higher prices. These are: (1) a Nash-Cournot model of competition and (2) a dynamic model that estimates a bound on consumer willingness-to-pay for DBS and thus indirectly measures incentives of New EchoStar to increase price. These are both in the category of a unilateral effects analysis because they explore the incentives of the merged firm to raise price under the assumption of no price response from cable suppliers.

Nash-Cournot Competition

On the assumption that EchoStar and DirecTV behave as Nash-Cournot competitors, the premerger prices for each network would satisfy the following condition:

\[
\frac{P_i - MC_i}{P_j} = s_i \eta
\]  

(2)

where \( i \) is firm /s share of DBS programming and \( \eta \) is the elasticity of demand for DBS expressed as a negative number. The right-hand side of equation (2) is the reciprocal of the firm-specific elasticity of demand for product \( j \). For example, if EchoStar has 40 percent of DBS subscribers and DirecTV has the remainder, and the elasticity of demand for DBS is \(-2.54\), then the firm-specific elasticity of demand for EchoStar would be about \(-6.4\), and the firm-specific elasticity of demand for DirecTV would be about \(-4.2\). These different firm-specific elasticities

\textit{Rozanski and Thompson (2005) discuss econometric issues addressed by the Antitrust Division in their investigation of the proposed merger.}

\textit{A Nash-Cournot competitor maximizes its profit by choosing its output level under the assumption that the outputs of other firms are fixed. This is not an accurate characterization of competitive conduct in an industry such as DBS, where suppliers set prices, and capacity constraints often are not binding. It is useful, however, as a baseline for premerger conduct that is neither especially intense nor collusive.}

\textit{The relevant share in equation (2) is the share of DBS, not the share of all MVPD services, because the market elasticity in equation (2) is the elasticity of demand for DBS.}

\textit{A more general model could incorporate perceived competitive responses from other firms. See, e.g., Baker and Bresnahan (1988).}
imply different prices, unless the networks’ marginal costs differ as well, which is plausible. The observed prices of $30.99 per month for EchoStar and $31.99 per month for DirecTV, plus $5.99 per month for local channels, are consistent with the Lerner Indices implied by equation (2) if DirecTV’s marginal cost is $28.94 and EchoStar’s marginal cost is $31.20. These estimated marginal costs are close to the levels estimated independently by MacAvoy (2002).

Under the assumption that the proposed merger would not generate efficiencies or create additional costs, a conservative assumption is that New EchoStar’s marginal production cost would equal the lower of the two estimated marginal costs from the Nash-Cournot model. This is $28.94 for DirecTV. Using this marginal cost and the DBS elasticity of 2.54 gives a postmerger price of $47.73 per month for DBS service; that is, from equation (1),

\[ P = \frac{MC}{1 + \frac{1}{\eta}} \]

which implies

\[ $47.73 = \frac{$28.94}{1 - \frac{1}{2.54}}. \]

The average of the premerger prices for EchoStar and DirecTV is about $37.48 per month. This is the simple average of the monthly prices for EchoStar and DirecTV, plus $5.99 per month for local channels. Under the Nash-Cournot assumption for premerger competition, the merger would cause a price increase of about 27 percent above premerger levels.

The predicted price increase from the merger depends on estimated price elasticities, the intensity of competition before the merger, and estimated marginal costs before and after the merger. The intensity of premerger competition affects the predicted price increase from the merger because the premerger price-cost margin is the basis for estimating the marginal cost of satellite TV, which in turn is an input for estimating the postmerger price. The estimated price increase from the merger would be larger than predicted with the Nash-Cournot model if premerger competition between EchoStar and DirecTV were more intense than implied by the Nash-Cournot assumption and lower if competition were less intense. In the former case, the estimated marginal cost would be closer to the premerger price, and equation (1) would then imply a higher price increase from the merger. If hypothetically, EchoStar and DirecTV were avoiding price competition premerger, then the estimated marginal cost would be lower, and the price increase from the merger would be smaller or even nonexistent.

Page 128
The parties argued that the merger would not raise prices in part because they claimed that consumers did not view EchoStar and DirecTV to be close substitutes and therefore the merger would not have eliminated significant price competition. According to the parties, competition was primarily between DBS and cable, not between EchoStar and DirecTV. The parties justified this conclusion using “churn data,” which estimated the fraction of subscribers leaving a particular DBS provider who went to the other DBS provider, cable, or over-the-air TV. The data showed that consumers switched between DBS and cable more than they switched between EchoStar and DirecTV.

Churn data, however, are difficult to interpret because consumer switching behavior is indicative of price elasticity only if the switching occurs as a direct result of a change in relative prices. Consumers switch MVPD providers for many different reasons, and significant changes in the relative prices of EchoStar and DirecTV have been rare. For example, if a consumer stops subscribing to EchoStar for a reason other than an increase in the price of EchoStar relative to the price of DirecTV, it is unlikely that the consumer would switch to DirecTV if the consumer considers the two DBS suppliers to be close substitutes and their relative prices have not changed. Thus, churn data that show more switching between DBS and cable than switching between EchoStar and DirecTV are consistent with aggressive competition between EchoStar and DirecTV if the switching is for reasons other than changes in the relative prices of the two DBS suppliers.

The parties also argued that the merger would not raise quality-adjusted prices because New EchoStar would have lower marginal production costs and supply improved products. Applying equation (1), the postmerger price is proportional to the postmerger marginal cost. If the postmerger marginal cost is smaller than the premerger marginal cost, the postmerger price might be no higher than the prices that existed before the merger. Furthermore, new and improved products offered by the merged firm would benefit consumers, although in ways that are not directly captured by equation (1). We address the parties’ efficiency claims below and demonstrate that the merger was not necessary for EchoStar and DirecTV to achieve many of the claimed efficiencies.

Dynamics of Consumer Adoption of Satellite TV

The history of customer adoption of satellite TV provides additional evidence for concern about a price increase from the merger. The price of DBS relative to cable has fallen since EchoStar first offered DBS service in 1996, largely because EchoStar and DirecTV have slashed the cost to consumers for equipment and installations. Early adopters of satellite TV revealed a higher willingness-to-pay for DBS relative to cable than did later adopters. Competition between EchoStar and DirecTV generated consumer surplus for these early adopters, as the quality-adjusted price of satellite TV net
of equipment and installation costs fell relative to cable. Over time, the merged company would have an increasing incentive to capture the willingness-to-pay of these consumers in its installed base.

Figure 4-3 shows three measures of the price of DBS relative to cable programming over the period February 1998 through December 2001. The monthly charge measure compares (1) the monthly price of EchoStar’s AT100 package, plus an installation charge amortized over two years, with (2) the average monthly charge for cable service, including equipment costs. The “per nationally delivered network” measure divides the monthly charges by the number of national networks provided. The price per quality-adjusted channel acknowledges that an incremental channel in programming packages with larger numbers of channels may not be worth

This analysis is related to observationally based price discrimination described in Fudenberg and Tirole (1998), where a consumer’s past purchase behavior reveals information about her willingness to pay.

As was explained earlier, the prices of EchoStar and DirecTV were very similar over this period. These monthly cable-service prices are interpolated from annual averages of the monthly charges for the basic service tier (BST) and the major cable programming service tier (CPST), plus equipment, compiled from the FCC’s series of reports on cable industry prices at [http://www.fcc.gov/mbs/criptg.html](http://www.fcc.gov/mbs/criptg.html).

These are the channels traditionally thought of as “cable channels” as distinct from local programming. See, e.g., note 11 in FCC (2003b).
as much to consumers as an incremental channel in a smaller package. This measure divides the monthly charge by a concave function of the number of national networks provided. All three relative prices declined between 40 to 55 percent over this period. This suggests that New EchoStar could raise prices by a significant amount before early adopters of DBS would switch to cable.

Two factors have contributed to maintaining relatively low DBS prices: competition between EchoStar and DirecTV, and competition between DBS and cable. Although the merger would eliminate competition between EchoStar and DirecTV, New EchoStar still would have an incentive to attract new customers from cable and keep existing customers from switching to cable. New DBS customers are likely to be more price-elastic than are old customers, in part because old customers have revealed a high willingness-to-pay for DBS by subscribing when the price of DBS was high relative to cable. Under the assumption that New EchoStar cannot charge significantly different prices to new and old customers, New EchoStar would have an incentive to set a low uniform price to build its subscriber base. However, as the installed base of DBS subscribers increases over time, the merged company’s incentives would tilt toward raising prices to exploit the installed base of DBS legacy subscribers.

The authors developed a simulation model that integrated estimates of elasticities for legacy subscribers, price-dependent growth rates for new subscribers, subscriber churn, and switching costs to explore New EchoStar’s incentives for a postmerger price increase. Over a wide range of plausible parameters the model predicted postmerger price increases for DBS service of more than 15 percent.

In this analysis, the time path of prices reveals that customers who purchased DBS services in the past had a high willingness-to-pay for DBS. New EchoStar could profitably exploit these customers’ high willingness-to-pay when the installed base of customers is large relative to the flow arrival rate of new customers. Switching costs are a further reason why New EchoStar could profitably raise prices when its installed base of customers is large relative to the flow arrival rate of new customers. Switching costs arise from sunk expenditures by consumers on installation and equipment, long-term purchase contracts, and the time and inconvenience of researching MVPD alternatives and having one installed.

Nielsen ratings were used to rank networks in descending order. Each network was given a normalized rating relative to the rating of the top-rated network. The cumulative normalized rating as a function of number of channels was estimated to be $0.347 + 2.218 \ln(n)$, where $n$ is the number of channels in a programming package.

Offsetting this effect is the fact that marginal cost is lower for legacy subscribers than for new subscribers because the subscriber acquisition cost for legacy subscribers has already been sunk. This was taken into account in the simulation model developed by the authors that estimated the incentives of New EchoStar to raise prices over time.

See Klemperer (1987) for a general discussion of the competitive significance of switching costs and Wise and Duwadi (2005) for a specific application to DBS demand.
National Pricing

Proponents of the merger argued that a commitment to national pricing would protect consumers and is credible because both companies had always priced their monthly service packages on a national basis. National pricing, however, merely averages the price increase from the merger across all consumers. Unless the elasticity of demand for DBS averaged across all consumers is very large, a commitment to national pricing would not eliminate concerns that the merger would lead to higher prices for all consumers.

To see why a commitment to national pricing would not protect all consumers, note that the total demand for DBS services \( (Q_{DBS}) \) may be written as

\[
Q_{DBS} = q_{DBS}^c + q_{DBS}^{nc}
\]

where \( p_{DBS} \) is the uniform price of DBS service, \( q_{DBS}^c \) is the demand in cabled areas and \( q_{DBS}^{nc} \) is the demand in areas without cable access. The elasticity of total demand for DBS with respect to a uniform price \( (r_{DBS}) \) is

\[
\frac{\Delta Q_{DBS}}{Q_{DBS}} = \frac{\Delta p_{DBS}}{p_{DBS}}
\]

where \( \epsilon_c \) and \( \epsilon_{nc} \) are the shares of demand and \( r_{DBS}^c \) and \( r_{DBS}^{nc} \) are the elasticities of demand in cabled and noncabled areas, respectively.

The estimates of the elasticity of demand for DBS by Goolsbee and Petrin (2004) and others are estimates of the total demand elasticity, \( r_{DBS} \). These estimates are low enough in magnitude to justify concerns that the merger would lead to an increase in national prices. To the extent that the elasticity of demand is lower in absolute value in areas without cable, New EchoStar would have an incentive to choose higher prices for customers in regions with no access to cable (or to digital cable) than in regions with cable access if it did not establish a single national price. The analysis leads to concerns about significant price increases from the merger for all groups of consumers; national pricing would simply average the price increases for all consumers.

ENTRY

If the competitive effects analysis identifies significant concerns that a merger may harm competition, the next step in the analysis following the DOJ/FTC Horizontal Merger Guidelines is an evaluation of the potential for entry. The scope for entry to remedy any adverse competitive effects from the proposed merger of EchoStar and DirecTV was limited because the availability of orbital satellite positions is constrained by regulation.

EchoStar and DirecTV occupied the only three orbital satellite slots that were capable of providing full-CONUS coverage for DBS broadcasts.
using technology available at the time of the proposed merger, located at 101°, 110°, and 119° W.L. The only slots not controlled by either EchoStar or DirecTV at the time of the proposed merger were at 148° W.L. and 61.5° W.L. If a new entrant offers DBS service using a satellite at either of these “wing” locations, the signal from the satellite would have difficulty reaching both coasts due to the curvature of the Earth. Full-CONUS coverage would require at least two satellites, which would greatly increase the cost of entry. Furthermore, a backup satellite would be required at each location in case of equipment failure. Thus, entry at the wing slots would be much more expensive than entry using a full-CONUS slot and probably would be impractical, even under the assumption that access to the wing slots would have been possible.

Technological advances have made it feasible to locate satellites closer together than nine degrees of longitude to serve the same geographic area without creating harmful interference with each other, and this has been done for satellites that broadcast to Europe (Spectrum Five LLC 2004, p. 3). Nevertheless, entry would be extremely difficult because a new DBS carrier would have to obtain regulatory approvals for entry at a "short-spaced” orbital location, likely over the objections of New EchoStar. Furthermore, a new competitor would have to invest billions to design, construct, and launch satellites, design and manufacture set-top boxes and dishes, negotiate contracts for programming (some of which was licensed exclusively to either EchoStar or DirecTV), and create a distribution and installation network for the satellite and ground technology necessary to provide a nationwide DBS service.

No new entry has occurred by a competitor offering nationwide DBS service since the launches of DirecTV and EchoStar. In April 2002, SES Americom filed a proposal to locate a new short-spaced satellite at 105.5° W.L., and in February 2005 Intelsat filed applications for new DBS satellites at four orbital locations; however, there had been no regulatory action on these proposals as of early 2007. Clearly, opportunities for entry are much more limited for potential new DBS carriers than is typical for new competitors in many other industries.

Along with the previously discussed commitment to national pricing, the parties proposed as a remedy for the competitive effects of the merger that they (1) divest or lease substantial transponder assets to Cablevision Systems Corporation, owner of R/L DBS Company (known as "Rainbow"), (2) joint venture with Cablevision in the provision of set-top boxes and the transmission of local programming (called "local-into-local" broadcasts), and (3) provide assistance to Cablevision in establishing retail outlets. Rainbow was the licensee of eleven frequencies at the easternmost United States, DBS orbital slot at 61.5° W.L. The parties claimed that the additional

“EchoStar and DirecTV also have filed proposals for new satellite locations. On June 23, 2006, the FCC issued a notice of proposed rulemaking to consider these requests (FCC 2006a).
assets and other assistance would enable Rainbow to be a nationwide DBS competitor, which was unlikely to occur without the assets and other assistance from the merged company.

It was difficult to see how the proposed remedy would create a new DBS competitor strong enough to replace competition lost by the merger and quickly enough to protect consumers from the loss of competition. The creation of a new nationwide DBS service is extremely expensive and, even if economically feasible, would be unlikely to occur within the two-year time horizon considered in the DOJ/FTC Merger Guidelines. The terms of assistance from the merged company would have to be monitored and enforced. Furthermore, the new company would have its feet in both the cable and DBS camps with Cablevision as its owner, which provided cable services in the New York-New England region.

Rainbow has not emerged as a significant DBS competitor and likely would not have emerged as a significant competitor even with the assistance promised by EchoStar/DirecTV. In 2005, EchoStar acquired the orbital locations controlled by Rainbow and has used them to offer additional high-definition and local programming (FCC 2005).

**EFFICIENCIES**

Proponents of the merger argued that it would create substantial efficiencies, particularly in the use of scarce radio spectrum. By eliminating duplicative EchoStar and DirecTV channels, the parties argued that New EchoStar would be able to broadcast local channels to more communities, offer additional high-definition content, and be a more effective competitor to cable than either company could be on its own (Willig 2001, p. 4). There is significant overlap in the satellite broadcasts by EchoStar and DirecTV. The two firms had duplicative national programming on 150 channels and duplicative local programming in 140-175 channels, for a total of 290-325 redundant channels.

The parties claimed that the merger would enable the merged firm to provide local network programming for one hundred DMAs, additional high-definition programming, expanded programming for Alaska and Hawaii, greater programming diversity, and near video-on-demand (EchoStar et al. 2001). At the time of the proposed merger, EchoStar and DirecTV beamed local network programming into about forty DMAs, representing about 76 percent of U.S. television households.

EchoStar and DirecTV are like twin pipes from the sky, both of which are filled with nearly identical content. The efficiency argument is that New EchoStar would need only one pipe to broadcast the existing content, freeing up the other pipe for additional content, more local channels, or more programming in high-definition format.
AFTERMATH AND REFLECTIONS

EchoStar and DirecTV claimed that the merger would achieve significant improvements in DBS spectrum utilization and enable the merged company to be a more formidable competitor to cable. Furthermore, the efficiency argument implied a belief that neither DBS company, on its own, could or would make the improvements necessary to match the programming content provided by cable systems.

In fact, both EchoStar and DirecTV have found ways to increase capacity and expand programming both in absolute terms and relative to the programming content offered by digital cable systems. EchoStar currently offers a high-definition programming package with twenty-six channels and boasts that it has "America's largest HD lineup" (EchoStar 2007). DirecTV offers eight channels and a number of regional sports networks in high definition (DirecTV 2007a). Both companies also offer pay-per-view, premium channels such as HBO and Showtime, and access to local broadcasts in high-definition format. In 2005, cable companies offered an average of 11.6 channels in high definition, which includes premium and local broadcast channels in addition to national networks, and the number of high-definition channels available on cable systems appears not to have increased substantially in the past two years.28 Cable companies, however, have made strides in other enhancements to their networks, such as video-on-demand and Internet access.

The parties argued that the merger would give them the capacity to provide local broadcasts to about one hundred DMAs, compared to the roughly forty DMAs served before the proposed merger. In fact, EchoStar and DirecTV have each surpassed their predictions of what they could achieve with the merger. By the end of 2006, EchoStar offered local broadcasts in more than 170 markets across the country, representing 96 percent of all U.S. television households (Ergen 2006). As of 2007, DirecTV offered local broadcasts in 143 markets across the country, reaching more than 94 percent of the nation's TV households (DirecTV 2007b).

Experience demonstrates that the merger was not necessary to achieve many of the benefits that the parties promised would occur if EchoStar and DirecTV combined their operations. Since EchoStar and DirecTV abandoned their proposed merger, both companies increased programming content by using gains in signal compression technology and by launching or acquiring additional satellites. For example, since January 1, 2002, EchoStar deployed four new satellites, and DirecTV deployed five. At $200-$300 million each, the new satellites cost upward of $2-3 billion. It is plausible that the merger would have eliminated the need for some or all

28According to one industry source, in 2007 most large cable TV systems offered about eight or nine HDTV stations, usually including HDTV versions of the major premium networks and a few other choices (CNET 2007).
of these new satellites, although to achieve the benefits of consolidation the parties would have had to invest considerable sums to replace the incompatible platforms of EchoStar and DirecTV with a single platform for the merged company.

The billion-dollar question is whether consumers would have been better off if EchoStar and DirecTV had been allowed to merge. The merger might have allowed the companies to economize on satellite launches, and while New EchoStar would be the direct beneficiary of any savings, some of the benefits could have flowed to consumers. However, the merger posed a considerable risk of higher prices from the exercise of market power. The analysis earlier in this chapter estimated that the merger could allow New EchoStar to increase prices by as much as $10 per month.32 Even if the merger increased prices by only $2 per month, this would increase consumer expenditures by more than $500 million per year for DBS service, which likely would exceed the plausible annualized efficiencies from the merger.33 Of course, the tradeoff would be more favorable to the proposed merger if the welfare measure included a contribution from producer profits.

Soon after EchoStar and DirecTV abandoned their proposed merger in response to objections from the DOJ and the FCC, News Corp acquired a controlling interest in DirecTV, which the FCC approved with nondiscrimination provisions in 2003 (FCC 2003a).34 It is conceivable that the News Corp purchase of equity in DirecTV in 2003 provided needed cash for DirecTV to upgrade its infrastructure, which would not have occurred in its absence. It is difficult to test this hypothesis because DirecTV does not disclose its financial results. However, EchoStar does report its financial performance, and the evidence suggests that cash flow was not a significant constraint on its investments. While EchoStar reported net negative income from 2000 to 2002, its reported net income was positive in every year from 2003 to 2006. In any event, unlike the horizontal combination of EchoStar

32The gap between actual cable and DBS prices appears to be large enough to permit such a price increase. In 2005, EchoStar service cost about $43 per month for about 100 channels and local broadcast networks, while the average cost of digital cable with expanded analog channels was about $56 per month (FCC 2006b). However, these numbers should be interpreted with caution, given differences in programming packages, installation costs, and equipment fees.

33This calculation is based on 26 million DBS subscribers as of June 2005 (FCC 2005). A price increase of $2 per month is about 5 percent of the average price for DBS service in 2005. On the assumption of a DBS own-price elasticity of -2.5, a 5 percent price increase would reduce the number of DBS subscribers to about 23 million.

34 MacAvoy (2002) estimated that the consumer harm from the merger would range from at least $120 million to $700 million per year solely for 2.5 million customers in rural areas who were not passed by cable.

35 There were media reports that News Corp was in negotiations with General Motors in 2001 to buy DirecTV. See, for example, Grover (2001), Bryer (2001), and Cohen (2001). News Corp’s efforts were preempted by EchoStar’s merger bid, although ironically News Corp eventually won the prize.
and DirecTV, the News Corp equity purchase in DirecTV was a primarily vertical transaction that did not raise the same level of competitive concerns as the proposed merger of EchoStar and DirecTV.

REFERENCES


U.S. Federal Communications Commission. Hearing Designation Order, Application of EchoStar Communications Corporation (a Nevada Corporation), General Motors Corporation, and Hughes Electronics Corporation (Delaware Corporations) (Transfers) and EchoStar Communications Corporation (a Delaware Corporation) (Transferee), CS Docket No. 01-348, October 18, 2002a.


Margaret E. Guerin-Calvert

INTRODUCTION

On October 4, 2002, after almost a year of review, the Federal Trade Commission (FTC), by a 3-to-2 vote, closed its investigations into two proposed transactions involving Royal Caribbean Cruises, Ltd. (RCC) and P&O Princess Cruises pic (Princess) and Carnival Corporation (Carnival) and Princess. The two transactions were under review simultaneously; one (between RCC and Princess) was a “friendly” transaction, while the other involved a hostile tender offer. The transactions were also reviewed by competition authorities outside of the United States, including the European Union (EU) and the United Kingdom (UK).

The transactions and the competitive review are of interest from many perspectives to antitrust practitioners:

- The extensive review and its findings highlight the complexity of mergers that are reviewed by the antitrust agencies and the importance of the facts of the specific industry to the ultimate analysis and decision-making. Among the key elements in the cruise matter were the relative “youth” of the industry, with rapid and recent expansion of the population of “cruisers,” and the business model of cruise ship pricing (taking into account high fixed costs, on-board revenues, perishable inventory, and thus the importance of filling ships).

*The author participated in the economic analyses of the Carnival-Princess transaction for counsel to Carnival. Her work on this matter was done in conjunction with Janusz Ordover and while the author was at Economists Incorporated as a Principal. This chapter relies on information and materials in the public record.
• The competitive effects analyses undertaken by both the reviewing agencies and the parties reflected an interesting application of many aspects of the Merger Guidelines, including product market, geographic market, coordinated effects, critical loss and diversion analysis, and entry and expansion.

• The analyses were informed by relatively straightforward empirical analyses rather than extensive econometrics or merger simulations. As is described in greater detail below, the empirical analyses of the merger involved assessments of booking patterns (e.g., the timing of when reservations for cruises were made and the rate at which cabins were filling up on a particular cruise) as well as price dispersion analyses.

• The specifics of the review and the key elements of the decision making at each of the reviewing agencies were detailed to an extent greater than is typical for cases that were not subject to challenge. In fact, the "cruises" decision has been touted by the FTC and others as reflecting an increased effort at transparency in the reviews of transactions, including those that are approved.

This chapter addresses these key themes and the process and economic analyses that led to the ultimate outcome, which was a decision by each of the three regulatory agencies not to challenge either of the transactions. Subsequent to these reviews, revised bidding ultimately led to the consummation of the Carnival-Princess transaction.¹

The next section provides background on the transactions and the key issues raised in the antitrust review. After that, the following section provides a detailed assessment of the economic analyses of the transaction.² The next-to-last section sets out the result of the antitrust review and follow-up, and the last provides conclusions and insights.

BACKGROUND OF THE CASE

The cruise merger investigations arose with competing bids for Princess in late 2001. In December 2001, Carnival made a hostile tender offer for the shares of Princess, which followed a November 2001 announcement

¹For interesting discussion of the issues involved in the decision, spelled out in greater detail due to the transparency of information, which prompted an active exchange of views on elements of the analyses, see Grimes and Kwoka (2003); Coleman and Simons (2004); Grimes (2004); and Coleman, Meyer, and Scheffman (2003).


This section of the chapter relies heavily on the information and analyses released by the FTC and presented in documents from the EU and the UK as well as by various analyses of the underlying issues.
of a proposed transaction between Princess and RCC via a dual listing structure. At that point in time, the three companies represented three of the four largest cruise lines in an industry with a substantial competitive fringe. Each company’s operations included multiple brands with a number of ships in areas or trades throughout the world, including the United States (e.g., Caribbean and Alaska), Europe (e.g., the Mediterranean), and Asia. Other cruise lines, including Star Cruises of Hong Kong with its Norwegian Cruise Lines or NCL brand, also operated a number of ships in many of the same areas.

Review Process and Competitive Issues

Review of the proposed transactions commenced relatively quickly before the U.S. antitrust agencies as well as their European counterparts. The U.S. review, according to the statements released by the FTC in 2002, focused principally on the possible competitive effects of the transaction in North America and on North American customers. This review included assessment of the routes such as the Caribbean and Alaska. The EU review of the Carnival-Princess transaction focused on the effects on European-based customers for cruises within Europe as well as outside, while the UK review of the RCC-Princess transaction focused on the possible effects of the transaction on UK customers (whether the cruise originated in a UK port or in some other port, including those that were accessed by air or other modes of transport by UK customers).

Issues that were central to the ultimate decisions included:

- whether “cruises” constituted a relevant market, or whether the relevant market was broader and included land-based vacations;
- whether the high concentration of the antitrust market that was finally determined as relevant by the FTC was sufficient to presume a competitive problem, or, if not, what was the burden of proof;
- whether the transaction was likely to foster coordination among cruise lines or allow for unilateral conduct that might be anticompetitive;


For example, Carnival operated as CCL, Holland America, Costa Cruise Lines, Cunard Line, Seabourn, and Windstar with a total of more than forty ships. Princess operated as Princess, P&O Cruises, Swan Hellenci, Aida Cruises, Arkona, and A’Rose, operated twenty ships, and offered river cruises. For more detail, see European Union Commission (2002, p. 2). RCC operated under two primary brands, Royal Caribbean International and Celebrity, with a total of twenty-two ships at the time of the initiation of the review in the UK. See UK Competition Commission (2002, p. 7).

• whether a restriction of output by diverting ships to other routes in order to raise prices on a particular “trade” would be profitable—an analysis that involved assessments of the magnitude of such diversion that would be required to “move the price needle,” the likely response of competitors, and the relative gains and losses in profits from the diversion;  
• whether the merger would be likely to enhance the ability of cruise lines to price discriminate among cruisers (e.g., the ability to target specific categories of cruisers).

The following section provides an overview and summary of the primary issues. Materials used in this review derive from filings or public documents as well as agency statements that reference the views of each of the merging parties, the FTC, the EU, the UK, competitors, travel agents, and third parties. For ease of exposition, the analyses and conclusions reached by the FTC (or addressed by the FTC staff in subsequent papers) are used as the framework, followed by relevant complementary or opposing perspectives of the other agencies, third parties, or other participants on specific issues.

**BASIC DEVELOPMENTS—APPLICATION OF THE MERGER GUIDELINES FRAMEWORK**

The following sets out the antitrust analysis of the proposed cruise mergers using the framework of the *Merger Guidelines*, which was the primary methodology for analysis and exposition of the decision used by the FTC. The European authorities followed similar analytical processes, with an emphasis on theories of dominance.

The analytics depended on the answers to the first-stage question: Were cruises a relevant market, or were they part of a larger “vacation” market? If the answer to the question was that the relevant market in which to assess the transaction included noncruise vacation alternatives, all parties agreed that the resulting shares of the combining firms in either transaction would be sufficiently low as to raise no substantive concerns. Alternatively, if cruises were the relevant market, then market shares and concentration

---

Third parties included travel agents, competing cruise lines, and the American Antitrust Institute. The EU document provides references to a number of submissions by various parties and provides some summary of the content of the various filings. See also third-party documents that are available separately, such as the AAI study and paper (Grimes 2002), and more detailed articles on the FTC analysis presented in a number of papers, articles, speeches, and PowerPoint presentations by FTC staff and officials. See, for example, Federal Trade Commission (2002).

Cruises have a geographic dimension that could cause market definition to be narrower than all cruises. For example, major destinations served by cruise lines from ports in North America include Alaska and the Caribbean, as well as South America and Canadian destinations from the East Coast.
levels could raise threshold concerns about the transaction and would require additional substantive analyses with regard to competitive effects as well as entry and expansion.

**Product Market Definition**

Analysis of product market definition comprised a substantial portion of the detailed statements of the antitrust authorities and also of the third-party submissions. The critical question was the extent to which land-based vacations constituted a sufficiently close substitute for cruises that a hypothetical monopolist of cruise lines would not be able profitably to sustain a small but significant and transitory increase in price (SSNIP). The SSNIP test was employed in the cruise merger, with varying estimates of elasticities presented in the statements or materials related to the cruise case. In addition, issues of price discrimination were assessed to determine, for example, whether there was a sufficiently large and discernible group of customers who could successfully be targeted for a price increase.

A starting point for the analyses was a detailed description of the product, as well as the method by which the product tended to be sold. Among the salient features of cruises that were considered were: (1) Considerable portions of cruisers were first-time cruisers and less frequent repeat business; (2) the attributes of the cruise ship, including the number and types of cabins (e.g., inside versus cabins with balconies) and the trends in the types of cabins as well as amenities; (3) the importance of brand and types of cruisers attracted to particular brands or types of ships or locations; (4) the on-board amenities and revenues from on-board activities; (5) the duration of the cruise; and (6) the method by which cruises were booked and sold. With regard to the last, the FTC noted that the data and information that the agency had reviewed supported the conclusion that the majority of cruises was booked through travel agents. In addition, each of the agencies noted the substantial period of time before sailing during which bookings were made (and cancelled), the fact that few customers paid "list" price, and the fact that the availability of substantial capacity near the sailing date tended to result in price reductions.

---

9 The SSNIP test is set out in U.S. Department of Justice and Federal Trade Commission (April 8, 1997), "Horizontal Merger Guidelines" at Section 1.0; there are numerous articles setting out the methodology for applying the SSNIP test generally, as well as in this matter. See, for example Verouden (2004) and Coleman, Meyer, and Scheffman (2003).

10 "Frequency of repeat" refers to both the tendency of customers to take additional cruises once they have cruised and the timing of their taking of additional cruises. Data discussed by the FTC indicated that customers tended to repeat on an infrequent basis, suggesting that cruise lines had to continue to attract new customers to cruising in order to fill current and expected future capacity.

11 "UK Competition Commission (2002, pp. 11-12). IRN Research (2002) examines whether in choosing their vacation type, cruisers had considered land-based vacations as well."
In assessing the market definition question, the FTC considered and evaluated the results of tests conducted on the elasticity of demand for cruising and the evidence related to the willingness and ability of customers to switch from cruises to land-based vacations in the event of a SSNIP. FTC economists (Coleman et al. 2003, p. 12) noted that the tests included examination of the results of a “natural market” experiment in which substantial capacity had been added to cruises within a relatively short period of time and that the capacity was filled without a substantial reduction in price. Based on analyses reported as confidential, the staff concluded that a 5 percent across-the-board price increase would not be profitable, since the demand elasticity was estimated to be about -2.0 and potentially higher, depending on the precise estimate of margins.

The EU cited similar analyses and referred to papers submitted to the Commission by NERA Economic Consulting, which conducted similar assessments of Princess pricing and profitability during the period of substantial capacity expansion, and a paper, “Profits and Prices in Evaluation of Capacity Expansion,” by Guerin-Calvert and Ordover submitted based on Carnival data. These citations and the discussion in the EU document provide insight into some of the analyses conducted by the parties’ economists and their perspectives on issues of market definition (as well as competitive effects). The two papers demonstrated that the substantial increases in capacity had been “absorbed” without significant diminution in prices or profits, indicating that the results were consistent with a broader market. The Commission expressed some concerns that there may have been some unqualified shifts in demand due to marketing or to premium cruise promotion that could affect the interpretation of the results and the conclusions with regard to market definition.

The FTC undertook or considered additional types of analyses. The Commission’s initial analyses had assumed an across-the-board price increase as part of the SSNIP test. Given the use of yield management systems by the cruise lines, the FTC hypothesized that the result of the SSNIP test could vary if one assumed that the hypothetical monopolist used yield management with some ability to target specific customers and thereby raise average prices. Based on this extension, the FTC concluded that while

---

12 The staff apparently also conducted critical loss analyses to quantify the extent of diversion required to discipline a 5 percent or 10 percent price increase.


14 The FTC papers provide detailed examination of various empirical analyses conducted by economists. The analyses conducted involved both cross-sectional and time series analyses. These included examination of average prices paid (per diems) for a given type of cabin at different points in the booking cycle for a specific cruise, as well as examination of the average prices paid for similar cruises (approximated by different sailings of the same ship) for a given type of cabin. See Coleman, Meyer, and Scheffman (2003).

the overall elasticities were high relative to critical loss estimates, the evidence "on balance" was consistent with a market that was narrower than cruises plus land-based vacations and should be limited to cruises. The European Commission undertook its own evaluation of the market definition issues, ultimately employing (if not declaring itself satisfied with) this perspective.

Finally, the FTC examined the possibility of price discrimination, including various types of customers who could be targeted as less price sensitive. Among the types assessed were repeat purchasers, "early bookers," and customers for certain types of "premium" cabins. Coleman et al. (2003, pp. 14-15) report that analyses showed that "there is very substantial variation in prices that cannot be explained by any characteristics of customers, transactions, ships, or trade. That is, the practice of yield management in the cruise line industry does not generate prices that can be adequately explained by customer or transaction characteristics, which would be a necessary (but not sufficient) condition for a viable price discrimination market."

Geographic Market Definition

Analysis of geographic market proceeded along different lines within the various agencies and participants. The FTC evaluated the alternatives available to North American customers and included cruises from a number of regions or trades, including Alaska and the Caribbean in the market.

16 See Federal Trade Commission (2002, Sec. A). The FTC took the high elasticities and factors with regard to price discrimination into consideration in its analyses of competitive effects. The dissenting commissioners supported the FTC's overall conclusion with regard to market definition.

17 The UK document indicated that there were mixed opinions as to whether cruises were part of a broader market and also indicated some support for the conclusion that cruises could potentially be "segmented" into more than one market based on the quality of ships and service. Nonetheless, the view was expressed that even if the market were narrower on either of these dimensions, the presence of the other qualities of cruises (e.g., higher or lower) and the presence of vacation alternatives served as some constraint on the companies' pricing. See UK Competition Commission (2002, p. 16). The EU also considered the issue of quality segments and concluded that there was insufficient evidence to establish separate markets, albeit noting that luxury cruises might constitute a relevant market. The EU also considered and rejected the concept that oceanic and river cruises were in the same market. The EU also excluded land-based resorts as a sufficient constraint on pricing to warrant inclusion in the market. See European Union Commission Decision, "Case No. COMP/M.3071" at p. 15. Both simultaneously with the FTC review and subsequently, papers have assessed the issue of market definition. In a letter to the FTC, for example, the AAI (Grimes 2002) expressed the view that examination of the customer profiles (e.g., many cruisers were older and were repeat customers), differences between land- and cruise-based travel, and the way in which cruise travel was sold suggested that the broader market was not supported.

18 The existence of other variation does not logically preclude the possibility of price discrimination; there is the prospect that the latter becomes harder to isolate in the presence of other factors.

19 The UK reached similar conclusions and focused on UK customers and concluded that the market included a broader set of cruises than those originating in or near the UK, noting that a large proportion of UK cruisers frequented Caribbean cruises. The EU focused on narrower markets
Market Share and Concentration

In the cruise market as defined by the FTC, market shares and HHIs were calculated based on lower berth capacities, and the mergers were found to result in a significant increase in concentration in a highly concentrated market. The dissenting commissioners emphasized that "If Royal Caribbean succeeds in its offer for Princess, the postmerger HHI will increase from a highly-concentrated 2800-plus to over 3700. If Carnival is successful in its bid for Princess, the HHI likewise will rise to close to 3800. Either merged entity would control nearly half of North American cruise capacity."

It should also be noted that in addition to the three companies involved in the transaction, other large players were Star (with about 11 percent share worldwide) and about 25 percent of capacity split among a number of other competitors. The EU and the UK also estimated shares of relevant markets with the UK focused on both a "premium" market and a broader market and concluded that the combined shares when taken into consideration with other characteristics did not indicate substantial concern about concentration. EU statistics showed high concentration in some of the affected markets.

The dissenting commissioners at the FTC expressed concern over the high levels of concentration and the increase in concentration from the merger. This concern was also stated in the letter to the FTC by the AAI (Grimes 2002), which raised the question of whether such a large increase in an already highly concentrated market should not be presumed anticompetitive.

Competitive Effects Analysis

The market definition and shares were the first stages of extensive empirical analysis of the likely competitive consequences of the proposed transaction.

---

Anthony and Thompson, FTC Dissenting Statement (2002). "If a North American cruise market exists, concentration increases by 1,848 to 4,474 HHI due to the Carnival/Princess combination" Moss (2002). Grimes and Kwoka (2003) criticize the FTC for not relying more on the presumption that high shares and concentration and substantial increases in concentration create market power issues. The EU considered a variety of share measures and expressed a preference for passenger cruise days.

These shares are reported in Coleman, Meyer, and Scheffman (2003, p. 9).

In addition, Grimes and Kwoka (2003) raise the issue of the presumption of high market share and concentration. The concept of presumption of higher shares is raised at times to emphasize that as a legal procedure this could shift the burden to the parties to disprove likely competitive harm, rather than requiring a demonstration of likely harm with some specificity by those in opposition to the merger.
In conducting its analyses, the FTC analyzed both the theory of "coordinated effects" and also of "unilateral effects." The latter theory involved situations in which a merger in, say, a differentiated products industry combined two close competitors so that each could now increase its own price unrestrained by its closest rival to whom customers might have switched.

In its investigation of unilateral effects, the FTC undertook various empirical analyses. These included analyses of the profitability associated with shifting capacity out of one area (e.g., North America) into another by the merged parties in anticipation of capturing sufficient volumes to be able to benefit from the shift. These analyses concluded that there was not sufficient predictability that would allow the parties to identify the amount of diversion between the two parties that would be sufficient to make unilateral price increases profitable.

Having rejected unilateral effects as an appropriate theory, the FTC considered coordinated effects theories both on pricing and on quantities.

**Coordinated Effects—Pricing**

As noted above, the coordinated effects theory of competitive effects analysis was regarded as the more relevant theory for this matter. The analysis of this possibility undertaken by the FTC was quite complex, given that an across-the-board increase was deemed not to be profitable. Instead, coordination would have to be selective. The FTC considered such factors as the transparency of price and price change information, the incentives of the merging parties and others to coordinate on pricing, and also the complexity of the overall price mechanism in the industry. Ultimately, the FTC analyzed the following matters:

- the complexity of prices, including the various dimensions on which price could vary;
- the complexity of any price discrimination scheme;
- the nature and extent of nonsystematic variation in prices based on head-to-head competition between similar cruises at similar points in time; and
- the variation in the prices in nonsystematic ways over the booking curve.

Ultimately, based on these analyses, the FTC concluded that the probability of price coordination was very low. In support of this conclusion, numerous graphical and tabular examples of the nonsystematic variation are included in the papers and presentations by the FTC economists.

---


24 The empirical analyses reported herein are those provided in publications by the FTC economists; extensive empirical analyses were conducted by consulting economists for the parties addressing similar economic theories but are not available publicly.
There are several interesting themes to these analyses. First, the analyses were extraordinarily data intensive. For example, one effort by the FTC’s economists examined the variability in average price per day paid by passengers for similar cruises. Figure 5-1, which was provided in a summary PowerPoint presentation by FTC economists, highlights the variability in pricing for comparable cruises. The analysis presented in the figure examines both booking levels and variation in price for a given type of cabin for a common itinerary.

The analysis is depicted over time, where the time is months prior to sailing. It illustrates the extensive variability in pricing over time and thus, arguably, the difficulty that the cruise lines would have in trying to detect “cheating” on any (implicit) coordinated effort to achieve higher prices.

Another of the analyses reported by the FTC economists involved an assessment of the average price paid per cabin for sailings on similar itineraries for all passengers in a given type of cabin. The analyses were conducted using a large number of cruises and a three-year time period. Figure 5-2 provides a representative example of the scope of data used in the analysis. The analysis depicted in Figure 5-2 involved determination of the "starting" or early price paid for three sailings of the same ship in three different years and an analysis of the price paid by all passengers in category 1 cabins on those sailings. The purpose of the analysis was to evaluate whether there was a common percentage discount from the starting price, which Figure 5-2 demonstrates was not the case.

*The analysis conducted uses a given type of cabin (e.g., an interior cabin or an exterior with a balcony) in order to be able to control for price differences that might exist due to quality differences.*
Second, this transaction investigated whether it would be difficult for competitors to differentiate "cheating" from competitive responses. This was a more recent development in empirical analyses of mergers at the agencies, and this was one of the first circumstances where this was expressly developed. The FTC concluded that it would be relatively difficult for competitors to identify cheating.

Third, the competitive issues were analyzed in the context of yield management systems. In the absence of systematic targeting of customers, yield management represented an interesting extension of the competitive effects analyses of yield management as it had developed in airlines.\footnote{Athey, Bagwell, and Sanchirico (1998).}

Finally, the staff expressly used "critical loss" analysis in market definition and competitive effects. This approach focuses on the likely amount of output sacrifice that is necessary for a firm that is interested in raising prices to see whether the action is rational under the demand circumstances in the industry.\footnote{See, e.g., Borenstein and Rose (1994).}

**Coordinated Effects—Capacity**

Having rejected the prospect of coordinated effects specifically on pricing, the FTC examined the likelihood that it would be profitable to reduce

\footnote{There are a number of articles addressing critical loss analysis that set out both the basic methodology and applications. These include Harris and Simons (1989); O'Brien and Wickelgren (2003); and Katz and Shapiro (2004). These articles set out the debate over whether critical loss analyses may understate the volume of sales that need to be diverted.}
capacity sufficiently so as to increase prices for cruises. The empirical analysis relied upon the hypothetical diversion of sufficient ships to other regions and assessment of the effects on overall margins from the shift. The conclusion was reached that the capacity reductions to increase prices in North America would have to be sufficiently large that it would not be profitable for the company as a whole, given the profits in the other regions and the effect of expanding capacity in those areas.

While the treatment of the capacity allocation and diversion theory was relatively terse in the FTC’s decision, this received far more extensive treatment in the subsequent articles. It was also the source of a sharp difference in view between the dissenting commissioners and the FTC majority. The dissenting FTC commissioners expressed the view that the companies monitored each other’s pricing and behavior and that postmerger concentration would be at a sufficiently high level to create the incentive to coordinate on pricing. The concern was raised that capacity additions could perhaps be delayed or reduced and thereby would provide greater opportunity for coordination and some reductions in capacity.

The EU and UK decisions as well as the FTC decision provide some answers on this point, including the conclusion that it would be prohibitively costly for all of the ships on order to be cancelled due to penalties that would have to be incurred with cancellation.

Entry and Expansion/Repositioning

Entry and expansion prospects focused in part on the ability of the companies to divert or shift cruise ships from one area to another but particularly focused on the number of new ships for which there were firm commitments with shipyards that could not be delayed or eliminated without substantial penalties. There was some difference in view as to the nature of contracts and the extent to which capacity was committed. The FTC concluded that there were substantial penalties associated with canceling contracts and that it would take considerable capacity reduction in order to increase prices.

OUTCOME AND SUBSEQUENT DEVELOPMENTS

All three reviewing agencies reached the same conclusion: approval of the cruise line mergers that were under their jurisdiction and review. As noted
in the foregoing, however, the bases for the conclusions and decisions were
different across the regulators. Each regulator expressed relatively similar
overall views with regard to the market definition issues in determining that
additional analyses of competitive effects were required to assess the com­
petitive effects. The EU as well as the UK focused considerably more atten­
tion on customer switching and customer choice of cruises as competitive
constraints and placed less weight on some of the empirical analyses. The
FTC relied heavily on empirical analyses to demonstrate that theories of
coordinated and unilateral effects were not sustainable, given the facts of
the industry.

The approaches developed by the agencies in this particular case
appear to have informed subsequent merger decisions and analytics as well
as the public announcements of such analysis. For example, the recent
Commentary on Merger Enforcement describes in some detail the specific
analyses conducted in cases and identifies the key issues in the matters.  

As noted above, the decisions indicated that the planned additions of
substantial capacity would provide a constraint on the merging parties and
that the merger would not likely result in reduced capacity or anticompeti­
tive effects. While there has been no detailed follow-up analysis provided
by the FTC, examination of the key trends in the industry shows that capac­
ity and pricing trends largely followed the expectations set out in the deci­
sion. The following sets out some of the key elements.

Examination of statistics compiled by the Cruise Line International
Association (CLIA), the cruise industry’s trade association, show that the
cruise industry continued to expand after the merger in terms of number of
passengers served. Figure 5.3 shows the overall increase in passengers
between 1987 and 2006 and depicts the destinations served as well.
Figure 5.3 shows that areas such as the Caribbean and Alaska continue to
grow, as do several other destinations—although it is difficult to determine
whether the expansion would have been larger (or smaller) in the absence
of the merger.

The growth in passengers on cruises is reflected in the underlying
growth in capacity. There were substantial additions of capacity after
2002-2003, as reflected in Table 5.1, which shows capacity additions
after 2003.

While there have been some increases in prices for cruises after 2003
based on an examination of company 10-K reports and trade press reports
that show some changes in prices in 2003 from 2002, these changes
appear consistent with overall increases in costs, particularly those associated

“Commentary on Merger Enforcement; see also Heyer (2005).

Also, shortly after the Carnival-Princess merger was approved, there were press reports that
expressed the concerns by rivals and travel agents about the expanded scope of operations of the
merged firm. See, e.g., Perez (2002).
TABLE 5-1
Growth in Number of Berths/Total Capacity: Operating Statistics of the North American Cruise Industry

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Ships</td>
<td>192</td>
<td>192</td>
<td>184</td>
<td>0.0%</td>
<td>4.3%</td>
<td>4.5%</td>
</tr>
<tr>
<td>Lower Berths</td>
<td>245,755</td>
<td>240,401</td>
<td>215,405</td>
<td>2.2%</td>
<td>11.6%</td>
<td>9.5%</td>
</tr>
<tr>
<td>Carrying (millions)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global Passengers</td>
<td>11.50</td>
<td>10.85</td>
<td>9.83</td>
<td>6.0%</td>
<td>10.3%</td>
<td>6.6%</td>
</tr>
<tr>
<td>Passengers Residing</td>
<td>9.06</td>
<td>8.31</td>
<td>7.48</td>
<td>9.0%</td>
<td>11.1%</td>
<td>7.0%</td>
</tr>
<tr>
<td>in United States</td>
<td>8.61</td>
<td>8.10</td>
<td>7.11</td>
<td>6.3%</td>
<td>13.9%</td>
<td>9.4%</td>
</tr>
</tbody>
</table>

with fuel. There is no indication in these sources of any anticompetitive changes in prices.”

CONCLUSION

The cruise mergers bring several elements of merger review into sharp focus. These include the importance of market definition for the initial competitive analyses of the transaction. For example, had the market been deemed to be broader by the reviewing agencies, then the mergers would have been approved substantially earlier in the investigations. In addition, the merger review included extensive exchange of data and information, as well as analyses, by a number of parties. While the review took ten months, it is clear that the development of data and information outside of the agencies was important to the overall timing and assessment, as were the development and analyses of information within the agencies.

One of the interesting observations that can be drawn from the merger review is the overriding importance of industry facts to the assessment and recommendations. The nascent stage of the industry, the fact that large numbers of ships were to be added and that diversion or reduction of capacity was very costly, the inability to articulate clear price discrimination markets, and the extreme variability of prices in a nonsystematic way—driven by many factors, including the relative elasticity of demand, high fixed costs, and perishable inventory—all came together to point to the conclusions that consumers would not be harmed by the transaction.

Review of the analyses conducted in the various jurisdictions is of value to antitrust practitioners because it highlights some similarities in merger review guidelines but some differences in approach. While all three agencies approved the merger, there were different perspectives on the scope of geographic markets (for example, whether to define relevant markets based on customers or on countries), on the importance of brand and smaller markets based on quality of product (and what constitutes a sufficient constraint on pricing between differentiated products), and on market definition and competitive effects.

But these are ultimately factual questions, and, as this chapter reflects, they were heavily debated throughout this case in its various forums. The release of certain empirical analyses, the detailed papers, and the substantive commentary provide insight into the framework employed and a better understanding as to how comparable analyses can be conducted for other transactions.

---

Footnote: For contemporaneous documents, see Business Research & Economic Advisors (2004) and Tsao (2003). Carnival Cruise 10-Ks for 2003, 2004, and 2006 show that average ticket prices increased by 4-5 percent in the last two years, which is similar to the 6 percent increase between 2002 and 2003. Price changes between 2003 and 2004 were higher (at 17 percent for average ticket prices), but this reflects changes in the mix of cruises and the combination of the two companies.
REFERENCES


Clark, Donald S. Letter to Janet McDavid, "Re: Carnival Corporation's Proposed Acquisition of P&O Princess Cruises, File No. 021-00411" (October 4, 2002).

Clark, Donald S. Letter to Richard J. Urowsky, "Re: Carnival Corporation's Proposed Acquisition of P&O Princess Cruises, and Royal Caribbean Cruises' Proposed Dual Listed Company Combination with P&O Princess Cruises, File No. 021-00411" (October 4, 2002).


CASE 6
Efficiencies and High Concentration:
Heinz Proposes to Acquire
Beech-Nut (2001)

Jonathan B. Baker*

INTRODUCTION

Mergers may benefit the economy by allowing firms to reduce costs or develop better products. Yet, the Supreme Court has in the past questioned whether cost savings or other efficiencies from merger should ever count in favor of a transaction that increases market concentration substantially. In 1967, the Court announced that "[p]ossible economies cannot be used as a defense to illegality" in merger review. The Court had previously explained that for courts to assess whether social or economic benefits can justify an otherwise anticompetitive merger goes beyond the ordinary judicial competence and is inconsistent with congressional intent.

The government's Merger Guidelines of the same era, issued in 1968, were only slightly more sympathetic: They recognized that efficiencies might justify an otherwise anticompetitive merger, but limited the defense to "extraordinary cases."

The case against considering efficiencies in the analysis of mergers among rivals (horizontal mergers) has, perhaps surprisingly, been made forcefully by both Richard Posner and Robert Bork—influential antitrust

The author testified as an economic expert on behalf of the merging firms. He is indebted to Michael Black, Steven Brenner, Christopher Cavanagh, Edward Henneberry, Richard Higgins, John Hilke, David Glasner, Richard Ludwig, Steven Salop, Marc Schildkraut, and John Woodbury.

commentators who have enthusiastically encouraged the courts to consider efficiencies in antitrust analysis generally (Posner 2001, pp. 133-143; Bork 1978, pp. 123-129). Posner and Bork contend that an efficiency defense would make merger analysis intractable in litigation, for several reasons. The cost savings likely to result from merger are easy to claim but may be hard to prove or disprove, particularly given that most of the relevant information is in the hands of the merging firms, who are interested parties. It may also be difficult to determine how soon the same efficiencies might have been realized through less restrictive means short of merger, such as internal growth or managerial changes. In addition, courts may be unable practically to weigh the efficiency benefits of a transaction against the likely harm to competition from loss of a seller. Accordingly, Posner and Bork conclude, the best institutional means of accommodating the concerns about the competitive harms from mergers among rivals in concentrated markets with the concerns about deterring efficient transactions is to address the possibility on average rather than case-by-case: by allowing mergers to proceed unless market concentration is substantial, and ignoring the efficiency justifications in specific cases otherwise.

Notwithstanding these concerns, the trend in antitrust law has been toward considering the efficiencies from merger in individual cases. Although the Supreme Court decisions from the 1960s that seemingly forbid an efficiency defense remain formally controlling, they are widely understood today as reflecting the perspective of an earlier era during which merger law was thought to vindicate noneconomic concerns, such as halting trends toward market concentration in their incipiency and protecting small business (along with the economic concern of preventing the exercise of market power). As such concerns have come to take a back seat to economic concerns across much of antitrust, judicial hostility to efficiencies has steadily decreased. Since 1979, the Supreme Court has recognized an efficiency defense in a related area of the law, the analysis of agreements among rivals (horizontal agreements) under the Sherman Act. Here, as elsewhere in antitrust, the Supreme Court has backed off from bright line rules in favor of flexible standards that permit judges to consider the full range of factors relevant to the analysis of likely competitive effects, and thus hold out the promise of reducing errors in deciding specific cases.

1 Other influential commentators, including Philip Areeda, Donald Turner, and Oliver Williamson, have been more sympathetic to allowing an efficiency defense to challenged mergers (Kolasky and Dick 2002; Williamson 1968).

2 Moreover, mergers may reduce profits on average (Ravenscraft and Scherer 1987), consistent with anecdotes suggesting that many claimed and even legitimately anticipated efficiencies never materialize.

3 Indeed, commentators do not agree how this weighing should occur in principle—particularly to what extent production cost savings not directly benefiting consumers of the products in the relevant market should count—as that issue implicates an ongoing debate over the purposes of antitrust law.

Taking their cue from this judicial trend, and perhaps also in recognition of the finance literature that emphasizes the importance of the “market for corporate control” (acquisitions) as a means of weeding out bad management and moving assets to their highest-valued uses, the federal antitrust enforcement agencies have been willing to consider seriously the efficiencies from merger in deciding whether to challenge proposed acquisitions for at least two decades. Similarly, the lower courts today do not interpret the 1960s Supreme Court decisions as foreclosing all consideration of efficiencies in the analysis of mergers.

While the modern trend in the lower courts is to recognize efficiencies as a defense, and the government’s Merger Guidelines were revised in 1997 to articulate a detailed approach to such a defense, the courts and antitrust enforcement agencies have nevertheless been skeptical about accepting an efficiency defense when market concentration is high. The Merger Guidelines state: “Efficiencies almost never justify a merger to monopoly or near-monopoly,” and efficiencies have never been the primary reason for the failure of a government challenge to a horizontal merger in court (Berry 1996, pp. 526-528; Conrath and Widnell 1999, pp. 688-690). Heinz’s proposed acquisition of Beech-Nut—a merger of two of the three leading U.S. producers of baby food—might have been the first court decision to uphold a challenged horizontal merger primarily on grounds of efficiencies, but a favorable district court decision was overruled by an appellate court panel unwilling to accept the efficiencies that had been credited by the district court judge in a market where postmerger concentration was very high.

Both sides in the litigation over the merger of Heinz and Beech-Nut accepted that the case turned on whether baby food consumers would benefit or be harmed by the proposed acquisition. As will be seen, Heinz and Beech-Nut claimed that the merger would produce variable cost savings that would lead to greater competition and lower prices within the relevant market, while the FTC contended that the merger would instead reduce efficiencies have, however, at times persuaded the antitrust enforcement agencies not to challenge a proposed merger among rivals.

This statement can be understood as premised in the economics of decision theory (Beckner and Salop 1999). Striking the right balance between adopting readily administrable decision rules and ensuring a full economic analysis is often difficult. Agencies and courts may thus be led to rely upon rebuttable presumptions based on easily observable information. From this perspective, the issue raised in Heinz is whether it is appropriate for courts and enforcers to depend on the shortcut of basing decisions primarily on market concentration when concentration is high, or whether a more complete competitive effects analysis would be practical and potentially lead to a different answer.

Thus, the case could be decided without reaching two contested questions on the welfare standard in antitrust law. First, when the merging firms participate in multiple markets, should courts permit a merger likely to raise price to consumers in one market on the ground that consumers in some other market would benefit even more through substantially lower prices? Second, should courts permit a merger likely to raise price on the ground that the production cost savings exceeds the allocative efficiency loss? This situation is most likely to arise when the merger would produce large fixed cost savings, unlikely to benefit consumers in the relevant market, while the output reduction within the relevant market resulting from the merger would be small.
competition, and prices would likely rise. Accordingly, the merger of Heinz and Beech-Nut provides an opportunity to consider whether efficiencies are an intractable subject for litigation and whether antitrust law properly accounts for the possibility that mergers among rivals could benefit society.

THE BABY FOOD INDUSTRY

In 2000, around the time of the proposed merger, three firms accounted for the sale of virtually all jarred baby food in the United States. Gerber, the leading firm, with more than 65 percent of U.S. sales, is a subsidiary of Swiss pharmaceutical conglomerate Novartis. Heinz, a food product company best known for its ketchup, had a share in excess of 17 percent. Heinz is also a leading seller of baby food in Europe. Beech-Nut, the major holding of Milnot, an investment firm, accounted for more than 15 percent of U.S. sales. Beech-Nut had previously been owned by seven different companies, including, most recently, two large food products concerns. Milnot acquired Beech-Nut in 1998 from Ralston Purina, which had purchased the company less than ten years before from Nestle. (Another firm, Earth’s Best, an organic brand, had a small and decreasing share, and was largely ignored during the merger litigation.)

Gerber and Beech-Nut are “premium” brands, with a strong reputation attractive to buyers. Heinz, in contrast, is a “value” brand, playing a role in the market similar to that of private label products in other markets. The Heinz brand name assures customers that the product is safe and healthy, but the brand appeals mainly to value-conscious shoppers who purchase it primarily because of its low retail price. At retail, Heinz sells at a discount averaging about 15 percent to the price of the Gerber and Beech-Nut products. Beech-Nut baby food generally sells for a price similar to that of Gerber’s products, or at a small discount.

Supermarkets do not carry all three major baby food brands. Shelf space and restocking costs are high for product lines characterized by many varieties sold in small jars, and evidently exceed the benefits of providing buyers with additional product variety once the supermarket has added a second brand. The great majority of supermarkets carry two of the three major brands; virtually none carries three, and a small fraction carries only one brand. There are no private label baby food products, presumably for a similar reason.

Gerber, the leading firm, is sold in virtually every supermarket. In contrast, Heinz and Beech-Nut compete to be the second brand on supermarket shelves. That competition is roughly even from a national perspective: Heinz is carried in grocery stores accounting for about 40 percent of supermarket sales, while Beech-Nut is found in stores with about 45 percent of grocery sales. But each second brand has areas of geographic strength: metropolitan areas where each is the second brand in most supermarkets and where the...
other brand has little presence. For example, Heinz is the primary second
brand in many cities in the Midwest, and Beech-Nut is the primary second
brand in many metropolitan areas in the Northeast and Far West. Both brands
have more than 10 percent of local sales in only about one-fifth of the major
cities, largely concentrated in the Southeast. It is not surprising that Beech-
Nut and Heinz exhibit strength in different regions: Scale economies in met­
ropolitan area distribution and promotion make it costly, for example, for
Heinz to serve a small number of retail outlets in a heavily Beech-Nut area.
Indeed, the firms do not generally consider themselves to be significant play­
ers in a city where they have less than 10 percent of local sales.

This pattern of distribution imposes transactions costs on shoppers
who wish to substitute between baby food products sold by Beech-Nut and
Heinz. The two firms’ products are virtually never on the same grocery
shelves. Moreover, a customer of Beech-Nut in a core Beech-Nut area will
typically have difficulty locating Heinz baby food products within the same
city, and vice versa.

The metropolitan areas in which both Heinz and Beech-Nut have a
nontrivial presence have probably been increasing as a result of supermar­
ket consolidation. Many chains prefer to carry the same brands in all their
stores, in order to obtain scale economies in inventory management, to
enhance their negotiating leverage in dealing with suppliers, and to guaran­
tee a consistent product line to those shoppers who patronize multiple
stores in the chain. (But not always: A supermarket chain may also wish to
vary some products across stores in order to cater to neighborhood tastes, as
with some ethnic foods.) On the rare occasions when a chain carrying
Heinz as its second baby food brand merges with a chain carrying Beech-
Nut, the merged firm may hold an “all-or-nothing” shelf space competition
between the two, ending with one brand taking over the other’s position on
the merged supermarket’s shelves. Shelf space competition also occurs in
other settings, as when Heinz or Beech-Nut approaches nonconsolidating
supermarkets with offers to displace its rival for the position on the shelf not
held by Gerber.

Beech-Nut and Heinz compete for the second slot on the supermarket
shelf and influence the retail price set by the grocery store through mer­
chandising payments from the manufacturer to the grocer (trade spending).
Most contracts between supermarkets and baby food manufacturers are not
fully specified in writing, so many of their aspects are worked out infor­
mally. The agreements typically involve a mix of fixed (lump sum) pay­
ments and payments that vary with the amount sold, roughly half fixed and
half variable on average. The fixed payments are payments for shelf space
for either new or existing products. The variable payments function more
like discounts from the wholesale list price and give the retailer an incentive
to lower the retail price: With a lower retail price, the grocer can sell more,
and thereby increase the payment from the baby food producer. Gerber, by
contrast, does not pay for or even compete for a place on the supermarket

Page 161
baby food shelf because its high market share and strong brand reputation automatically place it in a strong negotiating position with retailers.

THE PROPOSED MERGER

In February 2000, Heinz agreed to acquire the parent firm that owned Beech-Nut. The merging firms justified the transaction as a way to lower costs and improve products, much to the benefit of the companies and consumers. Heinz intended, after a year’s transition, to sell only under the Beech-Nut label, produce baby food exclusively in Heinz’s production facility, and share the cost savings from consolidating production with consumers by charging the low Heinz value price for the premium Beech-Nut product. The proposed merger would give Heinz distribution for its baby food in grocery stores accounting for the great majority of food sales.

Heinz planned to close Beech-Nut’s old, high-cost, labor-intensive production facility in Canajoharie, New York. That plant’s production lines are laid out vertically over several floors. Its production processes are relatively unautomated: They require frequent intervention by workers who must measure and add ingredients, move ingredients over large distances around the plant on carts, and set and monitor temperatures by hand. Production would be shifted to Heinz’s modern, highly automated Pittsburgh plant, which was reconstructed during the early 1990s as a facility for making baby food and other food products, such as private label soup. At Canajoharie, 320 workers produced 10 million cases of baby food each year, while 150 workers produced 12 million cases at Pittsburgh. Beech-Nut could not expand output cheaply at Canajoharie. In contrast, Heinz’s Pittsburgh plant was operating at only 40 percent of its dedicated baby food capacity and could add Beech-Nut’s production volume and still have 20 percent of the plant’s baby food capacity available for future growth. The merger would also allow the consolidated baby food production to take advantage of Heinz’s six regional distribution centers, which handle all Heinz’s food products, permitting the merged operations to share in the resulting scale and scope economies of distribution. The merging firms estimated that these efficiencies would permit Heinz to produce and distribute the Beech-Nut brand at a variable cost savings of about 15 percent.

THE FTC’s CHALLENGE

In July 2000, the Federal Trade Commission (FTC) decided to challenge Heinz’s proposed acquisition of Beech-Nut. That decision was controversial within the FTC. Both the legal and economic staffs that had investigated the merger recommended that the Commission decline to challenge it, and two of five commissioners voted against the challenge. The FTC’s
complaint seeking a preliminary injunction was tried in federal district court in late August and early September 2000.

The FTC emphasized that the baby food industry was highly concentrated, both nationally and in most metropolitan areas, and that the merger would increase concentration substantially. The FTC defined a jarred baby food product market and both a national geographic market and more localized city-specific geographic markets. Nationwide, the Herfindahl-Hirschman Index (HHI) of market concentration would rise as a result of the proposed transaction by 510 points to 5285, and in many cities the level and increase in concentration would be even greater. These are large numbers by the standards of the Horizontal Merger Guidelines or past merger cases. The number of significant sellers would fall from three to two. The Commission also observed, and the merging parties did not disagree, that entry would not be expected to defeat any exercise of postmerger Heinz-Beech-Nut’s market power. For the FTC, the transaction was in a highly suspect category: a merger leading to duopoly in a market protected by significant entry barriers.

To explain how competition would be harmed by this increase in market concentration, the FTC focused on the loss of wholesale competition between Beech-Nut and Heinz for shelf space. The FTC saw shelf space competition between the two as pervasive.” It argued that the repeated efforts that each firm made to take retail accounts from the other amounted to a constant threat to each other at all grocery chains. The loss of this wholesale competition for the second baby food slot on the grocery shelf would harm competition in several ways, according to the Commission.

First, the loss of wholesale competition as a result of the merger would remove a key impediment to tacit collusion among the baby food producers in their sales to consumers. Coordination between Heinz and Gerber is prevented today by the threat that Beech-Nut would respond to higher grocery prices for baby food by undercutting Heinz’s wholesale price to grocery stores and taking away Heinz’s shelf space. Similarly, Beech-Nut refrains from colluding tacitly with Gerber, even in core Beech-Nut cities where Heinz has limited presence, by the threat presented by shelf space competition from Heinz. Moreover, with a trend toward consolidation among supermarket chains, Heinz and Beech-Nut were increasingly obtaining significant space on retailer shelves in cities where that presence had previously been limited; this trend could only increase the competitive discipline conferred by shelf space competition between the two.

Second, the reduction in shelf space competition would give the merged firm a unilateral incentive to raise price without need for coordination with

“One FTC commissioner, in explaining his vote to sue, also argued that the merger would have been harmful even if competition between Heinz and Beech-Nut were instead limited, as the merging firms contended (Leary 2002, pp. 32-33). If the merging firms were not competing aggressively premerger, he suggested, that conduct would likely have reflected tacit coordination, which should not be rewarded by allowing the firms to merge and cease competing altogether. But the FTC did not allege premerger tacit coordination in court.
Gerber. With less shelf space competition would come less trade spending and thus less financial incentive for grocery stores to keep retail prices of baby food low. In addition, according to the Commission, Heinz’s plan to drop the Heinz brand in favor of the Beech-Nut label also would harm buyers by reducing consumer choice. In place of two alternatives to Gerber—a higher-priced Beech-Nut brand with strong reputation, and a Heinz brand that competed more on price than quality—consumers would have only one. Moreover, the merger would remove from the market a firm, Beech-Nut, that the FTC saw as innovative in the past, particularly in spurring the development and marketing of additive-free products.

Finally, the FTC regarded the shelf space competition at wholesale as worthy of protection under the antitrust laws for its own sake, independent of any effect it might have on retail prices or other dimensions of retailer conduct. Even if this competition did not benefit buyers directly, it would do so indirectly by benefitting grocery stores. Moreover, the FTC contended that harm to consumers should be presumed from the loss of wholesale competition in order to protect the government’s ability to prosecute substantial increases in concentration upstream, including mergers to monopoly, even if the effect of the transaction on retail competition is hard to identify.

With postmerger market concentration so high and entry unlikely, the FTC contended, the merger presented a clear and substantial danger to competition. Under such circumstances, efficiencies from merger could overcome the competitive concern only if the efficiencies were extraordinary indeed. But when the FTC reviewed the parties’ efficiency claims, those claims were found wanting. In particular, the FTC challenged the efficiency claims of the parties as not cognizable under Horizontal Merger Guidelines standards: They were unproved and overstated; they could have been achieved through practical means short of merger (e.g., investment in brand reputation by Heinz, plant modernization by Beech-Nut, or the sale of Beech-Nut to some other buyer); and they would result from an anti-competitive reduction in consumer choice (the loss of the Heinz brand). Even if these efficiencies were cognizable, the FTC also contended, they were insufficient in magnitude to outweigh the likely harm to competition, in part because they would not be spread across the entire output in the market where competition is threatened.

THE MERGING FIRMS’ DEFENSE

The merging firms saw Gerber as a dominant firm with the premerger ability to exercise market power, and Beech-Nut and Heinz as firms with

*In other industries, for example, this situation might arise if the upstream product (intermediate good) accounted for a small cost share of the downstream (final) product.*
limited competitive influence. Accordingly, the merger would do little to enhance Gerber's ability to exercise market power even absent efficiencies. Moreover, Beech-Nut and Heinz argued that their combination, by generating efficiencies, would promote competition with Gerber, leading prices to fall.

**Challenging Gerber's Dominance**

The parties' view of Gerber's role was based on evidence about market structure, firm conduct, and market performance. With respect to market structure, Gerber controlled more than 65 percent of baby food sales, and its only significant rivals, Beech-Nut and Heinz, had little incentive to challenge Gerber's dominance because they were each limited in their ability to expand. Beech-Nut could not increase output cheaply because it produced baby food in an old, high-cost plant. Its variable costs of manufacturing were 43 percent higher than those of Heinz (and likely also of Gerber), and its variable costs of production and distribution were 15 percent higher overall. Heinz could not expand cheaply because it sold a value brand that was limited in attractiveness to the many consumers who favored Gerber's and Beech-Nut's premium products.

Moreover, both firms were limited in their ability to expand, according to the merging parties, by the difficulty of obtaining distribution by grocery stores that did not already carry their brands. The second position on grocery shelves did at times change hands, but the shelf space rivalry between Heinz and Beech-Nut was circumscribed, and the incumbent firm held the upper hand. In an all-or-nothing competition to serve a grocery chain that previously sold Beech-Nut in some divisions and Heinz in others, the baby food manufacturer that already served most of the chain tended to win. Similarly, the firms' efforts to convince nonconsolidating supermarkets to displace their rival for the second position on the shelf had not led to much change in distribution patterns, as incumbents held an advantage in maintaining their existing shelf space.

The incumbency advantage had several sources. In order for Beech-Nut or Heinz to expand its distribution to stores it did not currently serve, it had to take shelf space away from the other contender for the second position on the supermarket baby food shelf. Doing so required it to outbid its rival for shelf space. In addition, the baby food producer had to compensate the grocery store for the costs of restocking stores and alienating the long-time customers of the rival brand. Moreover, the advertising and promotional costs of developing and maintaining a brand reputation are high. The payoff of such expenditures would be lower for Beech-Nut, for example, if it attempted to expand in a city in which it was carried on few grocery shelves, than it would be for Heinz, which had greater supermarket distribution in the same city. Under such circumstances, any promotional effort
that reached the entire city would largely go to customers who patronized stores where Beech-Nut’s product was not available.

The limited distribution of Beech-Nut and Heinz reduced the profitability of investments to develop significant innovation, forestalling these firms’ efforts to expand by introducing major new products. Heinz did not provide marketing and promotional support to products that were not available to at least 70 to 80 percent of customers, in part to avoid wasting a substantial fraction of its national advertising budget. In addition, access to less than half the potential market (40 percent for Heinz; 45 percent for Beech-Nut) limited these firms’ ability to spread the large fixed costs of new product development. Gerber’s incentive to innovate was also muted, in its case by the prospect that new products would primarily cannibalize its own dominant existing sales. Consistent with this view, aggregate sales of baby food had not been rising; retailers commonly described the category as “sleepy”; and most recent new product introductions were more marketing sizzle than substance, or else (as in the case of high-priced organic baby food) attractive to only a small segment of buyers.

Because so many grocery shelves were unavailable to Heinz, the company concluded before the merger that it would be unprofitable to introduce two major baby food innovations it had contemplated. The Environmental Oasis program was a “field to fork” quality assurance program intended to convince consumers that Heinz baby food is more nutritious and safe than anything they can make themselves. This program was successful for Heinz’s Italian affiliate, perhaps because of concerns about the food supply there in the wake of Chernobyl. Heinz also found that it would be unprofitable to introduce baby food produced using a new technology, aseptic production, that it had considered marketing as a high-priced, higher-quality product line for sale in addition to its regular brand. The aseptic production method would improve baby food taste by allowing Heinz to sterilize the product with less cooking time, and it would allow Heinz to introduce an attractive new product packaging (a microwavable and resealable pudding pack).

Firm conduct and market performance also suggest that Gerber was a dominant firm. Gerber was the pricing leader when wholesale prices changed; it arguably set the pricing umbrella for the industry. In addition, Gerber’s prices had been rising faster than the prices of food in general during a period in which input costs were largely unchanged, which was again consistent with some ability to exercise market power.

Absent the merger, the parties contended, neither Heinz nor Beech-Nut had much incentive or ability to take on Gerber. Without this merger, the industry would continue to perform poorly, with prices in excess of competitive levels and limited new product introductions. In contrast, the acquisition would promote industry competition by removing all the impediments to expansion by Gerber’s new (merged) rival. The merged firm could combine Heinz’s low-cost automated production process with Beech-Nut’s
premium brand reputation, allowing the merged firm to produce a premium product at less cost and making it profitable for that company to reduce the premium product’s price. Access to at least 85 percent of grocery shelves would make advertising and promotion cost effective in all major cities and would make it profitable for the merged firm to introduce innovations such as Oasis and aseptic production.

The high market concentration resulting from the merger was beside the point. It was not generated by the dominant firm’s entrenching its position by gobbling up a smaller rival; it flowed instead from the merger of two smaller firms that would create competition where competition had previously been limited.

Questioning the FTC’s Theories

In addition to explaining why the merger would promote competition, the merging parties also questioned the competitive effects theories proffered by the FTC. In doing so, they challenged both the FTC’s claim that competition between Heinz and Beech-Nut kept prices low premerger and the FTC’s view that the efficiencies from merger would have little competitive significance.

Unilateral Competitive Effects

First, the merging parties contended that the merged firm would be unlikely to raise the retail price of baby food unilaterally. In part, the parties argued, it was hard to be concerned about the loss of head-to-head competition for consumers when there was little retail competition between Beech-Nut and Heinz even before the merger. After all, the two brands were never on the same grocery shelf, and in most cities only one of the brands was widely available. Moreover, the merging firms and a number of grocers indicated that both Beech-Nut and Heinz priced against Gerber, not against each other.

The lack of significant retail competition between Heinz and Beech-Nut was also demonstrated using systematic empirical evidence. The merging firms presented estimates of cross-price elasticities of demand in a sample limited to metropolitan areas in which both Heinz and Beech-Nut accounted for at least 10 percent of retail sales (“mixed” markets), and thus both had a nontrivial presence in (different) stores. The study estimated that a 5 percent decline in the Beech-Nut price would lower the quantity of Heinz sold by about 0.1 percent (cross-price elasticity of demand of about 0.02)—a buyer response that is trivial in practical economic significance and insignificant statistically. By comparison, a 5 percent decline in the Gerber price would lead to a 3.1 percent reduction in the quantity of Heinz sold (cross-elasticity slightly greater than 0.6). Similarly, a 5 percent decline in the Heinz price would lead to a 0.6 percent
A 5 percent Gerber price decline would again have generated a larger output reduction, here 4.0 percent (cross-elasticity of 0.8). The FTC chose not to present alternative empirical demand elasticity estimates. Rather, it challenged the implications of these studies with grocery testimony that the two products competed at retail. In addition, the FTC raised a statistical concern about the merging parties’ demand elasticity study. It pointed out that the demand analysis was based on shelf prices, not transaction prices, because it did not account for the effects of promotional coupons issued by Beech-Nut and Heinz. The merging firms responded that the absolute level of merging firm couponing was small during the sample period and that the key cross-elasticity results were not biased from the omission of a variable accounting for this activity because almost all the coupons distributed by Heinz and Beech-Nut during the sample period were through direct mail, sent to new mothers in a steady, uniform way, uncorrelated with the retail price.12

The FTC again chose not to present a competing analysis of the data. Instead, it responded primarily with testimony from some grocers that where both are present in the same areas, they depress each other’s prices as well as those of Gerber. The merging firms also questioned the FTC’s claim that the loss of shelf space competition between Beech-Nut and Heinz would lead to higher retail prices for baby food. According to the firms, the incremental trade spending prompted by a bidding war for a retail account was mostly fixed, not variable. Grocery witnesses agreed with the prediction of economic theory: Increases in fixed trade spending would not be passed through to baby food buyers. The increased revenues could instead go to shareholders in the form of higher profits. Or, if those higher supermarket profits were competed away, the beneficiaries would be supermarket

---

12 The omission of an explanatory (right-hand side) variable does not bias the estimated coefficients in a linear regression if the omitted variable is not correlated with the included variables (Greene 2000, pp. 334-337).
customers generally—the store might widen aisles or stay open longer hours, for example—rather than baby food purchasers in particular.

The results of a third systematic empirical study presented by the merging firms were consistent with this economic analysis. The study analyzed the retail price effects of bidding competition for shelf space between Beech-Nut and Heinz. Grocery stores served by Heinz were divided into two groups, based on whether Heinz and Beech-Nut had competed for the shelf space at the time that the most recent agreement had been reached. The results showed no significant difference, either statistically or in practical economic terms, in retail prices for either firm with products on the shelf (Gerber or Heinz) related to whether there had been wholesale competition between Beech-Nut and Heinz in the past. In response, the FTC pointed to instances, drawn from grocer testimony or merging party documents, where the threat that a supermarket would switch its choice of second brand seemed to be associated with retail price competition.

Coordinated Competitive Effects

The merging firms questioned the FTC’s view that the acquisition would make tacit collusion among the major baby food manufacturers more likely. They questioned the feasibility of coordination in general, given the time lag in the ability of the two firms to detect wholesale price cuts by each other. They also cited a recent decision by an appellate court rejecting an allegation that the baby food producers had expressly colluded.

More importantly, the efficiencies that the parties saw flowing from the transaction played a central role in the merging firms’ response to the government’s coordination story. According to the firms, the efficiencies should not be understood as somehow justifying an otherwise anticompetitive merger; rather, they were the reason that the merger would promote competition affirmatively.

The cost savings from plant consolidation were, in the view of the merging firms’ efficiency expert (later quoted with approval by the district court judge), “extraordinary.” The variable costs of manufacturing the Beech-Nut line would fall by 43 percent, and the variable costs of all production and distribution activities would decline by 15 percent. The merging parties argued that these extraordinary variable cost savings would give the merged firm a powerful incentive to lower price. It would do better by lowering price and taking market share away from Gerber than by colluding tacitly with Gerber at a higher market price but lower market share.

Put differently, the merging firms argued that they would pass through all of the variable cost savings resulting from the acquisition to consumers. Heinz indicated that it would charge the low Heinz value price for the premium Beech-Nut product, thus lowering the price of Beech-Nut by about

\[\text{In re Baby Food Antitrust Litigation, 166F.3d 112 (3dCir 1999).}\]
15 percent, the full amount of the variable cost savings in production. This prediction was credible, according to Heinz, because of the company’s history of passing through similar cost savings achieved in the production or distribution of other food products in its portfolio, including cat food and ketchup. More importantly, an economic analysis showed that the likely pass-through rate in the absence of postmerger coordination was very high, at least 50 percent and quite possibly the 100 percent that the company claimed.

The pass-through rate is the proportion of variable cost savings that results in lower consumer prices. For example, if a firm lowered price by 5 percent when its variable costs fell by 10 percent, its cost pass-through rate would be 50 percent. Even a monopolist has an incentive to lower price if its variable costs decline: With a uniformly lower marginal cost curve and downward sloping demand, it necessarily equates marginal revenue with marginal cost at a higher level of output and lower price. For example, a monopolist with a linear demand will pass through 50 percent of any reduction in marginal cost."

The shape of the firm’s demand curve is an important determinant of the pass-through rate. The pass-through rate depends on the curvature of demand (the rate of change in the slope or elasticity). If a firm sells at a price in excess of marginal cost, as is common for producers of differentiated products, and the seller’s demand grows substantially more elastic as it lowers price, then the firm has a powerful incentive to give buyers the benefit of a cost reduction by reducing price. Under such circumstances, the seller would likely benefit more by lowering price to expand output, thus capturing a positive price-cost margin on the many additional sales it makes, than it would by keeping price (and thus the price-cost margin) higher while limiting its output expansion. The more that the firm’s demand curve is shaped in this way, the greater is the incentive for the seller to pass through cost reductions to buyers in the form of lower prices.

According to the merging firms, the demand curves facing both Heinz and Beech-Nut had a curvature that strongly favored a high pass-through rate. Heinz’s marketing experience was that its sales became dramatically more responsive to price cuts as the gap between its retail price and that charged by Gerber increased. Econometric estimates of the demand for both the Heinz brand and the Beech-Nut brand, based on functional forms that did not constrain the curvature of demand, also showed that each firm’s demand grew much more elastic when the firm’s price declined slightly—so much so as to make plausible Heinz’s claim that a cost reduction would be fully passed through to consumers. Rivals’ anticipated reactions to price changes can also affect the incentives to pass through cost reductions, but

---

"A firm’s pass-through rate can be interpreted as the ratio of the slope of firm-specific (residual) demand to the slope of firm marginal revenue. The intuition is that when faced with a reduction in marginal cost, the firm will increase output so that marginal revenue declines by the same amount. (Bulow and Pfeiderer 1983)."
simulation studies suggested that the competitive interaction among firms was a minor determinant of the pass-through rate for Heinz and Beech-Nut relative to the curvature of demand. Accordingly, this economic analysis demonstrated that the pass-through rate for cost savings from the merger of Beech-Nut and Heinz would likely be very high, at least 50 percent and quite possibly the 100 percent that the company claimed.

Heinz would want to compete aggressively postmerger by passing through all its baby food cost savings to consumers, according to the merging firms. The merger solved the problems that had previously limited both Heinz's and Beech-Nut's abilities to expand. It would give Heinz a premium brand, permit the Beech-Nut brand to be produced in a low-cost facility, and give the merged firm the national distribution required to make major innovation profitable. The transaction would thus effectively create a "maverick" firm with the ability and incentive to expand output (Baker 2002). Such a firm would not find it profitable to settle for its premerger market share, even if coordination with Gerber would facilitate an increase in the market price. Rather, with its new ability to expand inexpensively and to introduce new products, the merged firm would do better by taking market share away from Gerber. The resulting divergence in incentives between Gerber and Heinz would undermine the possibility of postmerger coordination. Not surprisingly, Gerber's internal documents predicted more intense competition following the merger; they did not predict a more cooperative environment.

If the merger led Heinz to compete aggressively and to pass through all the variable cost savings in production and distribution to consumers, then the quality-adjusted price of baby food would decline by 15 percent. Beech-Nut buyers would be able to purchase their premium brand for 15 percent less; Heinz buyers would be able to purchase a brand worth up to 15 percent more for the same low price they had previously paid; and Gerber customers would get the option to switch to a premium brand at a 15 percent discount, possibly leading Gerber to reduce price as well. In this way, the merging firms argued, the efficiencies would benefit all buyers in the baby food market.

**Loss of Wholesale Competition**

Finally, the merging firms questioned the FTC's view that the transaction should be blocked because it would end head-to-head wholesale competition for shelf space between Beech-Nut and Heinz, without regard to

---

15 The Heinz brand name would be withdrawn from the market, leading the FTC to voice a concern that consumers would be harmed from the reduction in product variety. The merging firms responded that the Heinz brand name primarily conveyed a low price to purchasers, not other characteristics valued by baby food buyers. In consequence, consumers would not be harmed by the removal of this brand so long as some other brand, here Beech-Nut, was sold at the low Heinz value price. Moreover, the merging parties noted, the merger would also increase product variety by facilitating the introduction of a new product line, based on Heinz's aseptic production process.
whether the FTC could demonstrate an adverse effect on retail prices. The loss of wholesale competition between the merging firms does not necessarily mean a loss of shelf space competition taken as a whole, because a reinvigorated Heinz would likely compete with Gerber for prime shelf space postmerger, inducing Gerber to pay grocers for shelf space for the first time. Moreover, any loss of wholesale competition is inextricably linked to the benefits that the merger creates for retail competition. The merging firms questioned whether a court should object to an increase in wholesale prices if that outcome were necessary in order to reduce the retail price of baby food.

THE COURTS DECIDE

The FTC, unconvinced by the merging parties’ arguments, voted to seek a preliminary injunction to stop the merger on July 7, 2000. A five-day evidentiary hearing in federal district court took place during late August and early September. The district court decision was issued on October 19, 2000.

The district court sided with the merging firms.16 The court agreed with the FTC that the high and increasing market concentration resulting from the transaction created a presumption that the merger would harm competition and that entry was not easy. But it found that the defendants had successfully rebutted the presumption arising out of market concentration by proving extraordinary efficiencies. "When the efficiencies of the merger are combined with the new platform for product innovation, . . . it appears more likely than not that Gerber’s own predictions of more intense competition . . . will come true."17

The district court rejected the FTC’s competitive effects theories. The merging firm would not be likely to raise the retail price unilaterally because retail competition between Heinz and Beech-Nut was limited and because the wholesale competition between the two did not benefit consumers. The court relied in part on the merging firms’ econometric evidence in reaching this conclusion and rejected the FTC’s claim that the exclusion of coupons from the price data made the studies unreliable.” Nor would coordination be likely in the market postmerger. Rather, the district court found it "more probable than not” that the merger "will actually increase competition” in the baby food market.”

17Heinz, 116 F. Supp. 2d at 199.
18Heinz, 116 F. Supp. 2d at 196 n.6.
The FTC appealed. On April 27, 2001, a unanimous appellate panel reversed the district court. The appeals court concluded that the efficiencies evidence accepted by the district court was insufficient, both as a defense and as a basis for showing that postmerger coordination would be unlikely. Without the efficiencies evidence, defendants could not prevail over the inference of harm to competition arising from the reduction in the number of sellers and the increase in market concentration.

The appeals court pointed out three main problems with the district court’s factual findings on efficiencies. First, the district court should have considered the reduction in total variable cost, rather than merely the reduction in the variable costs of manufacturing. Second, the district court should have analyzed the magnitude of the cost reductions over the merged firms’ combined output, rather than with respect to Beech-Nut alone. Third, the district court did not satisfactorily explain why the efficiencies could not be achieved through reasonable and practical alternative means, with less competitive risk than would arise from merger. The court of appeals suggested in particular that Heinz could have gotten to the same place by investing the money it was spending to acquire Beech-Nut on improving recipes and promoting a premium brand name.

The court of appeals also dismissed the district court’s conclusion that postmerger collusion was unlikely on the ground that defendants had failed to show that the difficulties of solving the “cartel problems” of reaching a consensus and deterring cheating “are so much greater in the baby food industry than in other industries that they rebut the normal presumption” of anticompetitive effect that would apply in reviewing a “merger to duopoly.” In addition, the appellate court found that the district court had erred in concluding that Heinz and Beech-Nut did not really compete at retail and in concluding that the merger would promote innovation.

The court of appeals went to unusual lengths to reverse the district court opinion. An appeals court must accept the district court’s findings of fact, unless the district court committed clear error. This highly deferential standard promotes the efficient use of judicial resources by limiting the scope of appeals. But here, the appellate panel engaged in what one commentator has termed “an extraordinary amount of appellate factfinding” (Kolasky 2001, p. 82).

“Judge Bork, whose influential antitrust book argues against permitting an efficiency defense to mergers, filed an amicus brief to the appeals court in support of Heinz’s acquisition of Beech-Nut. He contended that courts should allow a merger between two smaller firms that would create a stronger competitor against a firm that was dominant before a merger and would otherwise likely remain dominant. The brief did not address whether this position was tantamount to asking the court to accept an efficiency defense.


*Heinz*, 246 F.3d at 380-381.

“Compare, for example, the appellate court’s deference to the factual findings of the district court in *U.S. v. Microsoft Corp.*, 253 F.3d 5 (D.C. Cir. 2001), handed down by the same court of appeals a few months after the decision in *Heinz*.”
In concluding that the district court did not look at total variable cost or at cost savings spread over the merging firms’ entire output, the appeals court overlooked the evidence that Beech-Nut’s total variable costs of production would decline 15 percent by shifting production of its brand to the Heinz facility, and the evidence that Beech-Nut customers would pay 15 percent less while Heinz customers would obtain a premium product at a value price. The appellate panel’s theory that Heinz could have created a premium brand on its own had not been pressed by the FTC and overlooked the evidence that Heinz could not profitably make major new product investments absent national distribution, and had not done so. The appeals court’s conclusion that no reasonable district court judge could have found that Heinz and Beech-Nut did not really compete at retail relied on the anecdotal testimony of some witnesses that was inconsistent with the anecdotal testimony of other witnesses, and gave no weight to the systematic empirical studies of retail competition introduced by defendants and relied upon by the district court. In contrast, when the appeals court rejected the district court’s conclusion that the merger would promote innovation, its central complaint was that the evidence on which the district court relied was not statistically significant, and hence highly speculative.

The appellate court’s thoroughgoing rejection of the district court’s opinion appeared to be rooted in its skepticism about the efficiency defense, particularly when the merger would lead to a highly concentrated market. This concern was signaled in an order that the appeals court issued to prevent consummation of the merger while the appeal was pending, where the court noted: “[A]lthough there is much to be said for recognizing an efficiencies defense in principle, the high concentration levels present in this case complicate the determination of whether it should be permitted here.” The same theme was emphasized when the court of appeals issued its decision on the merits: “[T]he high market concentration levels present in this case require, in rebuttal, proof of extraordinary efficiencies, which the appellees [merging firms] failed to supply.” The court of appeals was so concerned about the competitive dangers flowing from high market concentration that it arguably created a new legal standard, raising the bar for the defense. In order for the district court to conclude that coordination would be unlikely in this highly concentrated market with entry barriers, the court of appeals held, the district court would have to find that the difficulties of colluding are not just high in some absolute sense; they must be “unique” to the baby food industry, and “so much greater . . . than in other industries that they rebut the normal presumption” that concentrated markets protected from entry are ripe for tacit collusion. The appeals court’s

---

24 Federal Trade Commission v. H.J. Heinz Co., 2000 Trade Cas. (CCH) P73,090 (D.C. Cir. 2000). Only one member of the circuit court panel issuing this order was on the panel that ultimately decided the appeal.

25 Heinz, 246 F.3d at 720.

26 Heinz, 246 F.3d at 724-725.
skepticism about an efficiencies defense when market concentration is very high stands in contrast to the way the courts treat entry: Ease of entry undermines proof of anticompetitive effect no matter how high the market concentration.\textsuperscript{27}

In defense of its perspective on high concentration and efficiencies, the appeals court speculated that even if the defendants were correct in predicting that the postmerger Heinz would compete aggressively with Gerber in order to increase its market share, that incentive might at some time dissipate, and the two would eventually come to see their interest in tacitly colluding rather than competing.\textsuperscript{28} Had it accepted the defendants’ view of premerger industry conduct, however, the court might have seen the possibility of increased competition followed by tacit collusion as superior to the present situation, in which the dominant firm, Gerber, continues to exercise market power free from serious challenge.

**LOOKING FORWARD**

The appellate court decision ended Heinz’s attempt to acquire Beech-Nut. This outcome did not arise as a matter of law: The legal result was merely that the district court was ordered to enter a preliminary injunction, leaving the parties free to pursue the case in a full administrative trial at the FTC. Instead, as is common in merger litigation, the delay that further proceedings would impose, combined with their uncertain outcome, undermined the business reasons for the transaction.

With respect to merger policy generally, the appeals court’s decision leaves open the possibility that an efficiencies defense could succeed in other cases—or even could have prevailed in this very case if the litigation were to have continued through an administrative trial.\textsuperscript{29} But the strong rejection of the “extraordinary” efficiency claims in Heinz calls into question the extent to which the courts will be willing to accept an efficiency defense when the market is highly concentrated.

**POSTSCRIPT**

The baby food industry structure has not changed much since the court case, although all three leading brands have new owners. Organic baby

\textsuperscript{27}U.S. v. Waste Management, Inc., 743 F.2d 976 (2d Cir. 1984); U.S. v. Baker Hughes, Inc., 908 F.2d 981 (D.C. Cir. 1990). It may be appropriate for courts to be more skeptical of an efficiencies defense than an entry defense, however, to the extent the evidence relevant to efficiency analysis is more likely to be under the defendants’ control and to the extent efficiency claims tend to be more speculative.

\textsuperscript{28}See Heinz, 246 F.3d at 725.

\textsuperscript{29}Heinz, 246 F.3d at 725.
foods have grown more popular, and new brands have entered into this niche, but their total market share remains small. The two major brands that sought to merge in 2001 have changed hands: Heinz sold its U.S. baby food division in 2002 to Del Monte, a large producer of canned fruits and vegetables; and Beech-Nut was sold in 2006 to Hero AG, a Swiss firm with other infant feeding businesses, mainly in western Europe. Also, Nestle agreed to buy Gerber in 2007.

Gerber remains the industry leader, and its market share has reportedly grown substantially, to about 80 percent. Gerber’s seeming entrenchment is consistent with the merging firms’ view that Gerber’s rivals posed only a limited constraint on the dominant firm’s ability to exercise market power. But the FTC would no doubt reply that Gerber’s current position is beside the point; the question in the case was whether the competition that Heinz and Beech-Nut posed for Gerber, however limited, would have been lost as a result of the merger.

Nor has there been much change in the receptivity of courts to efficiency claims in merger cases. In a merger case brought by the Justice Department in 2004, the merging firms’ efficiency claims were dismissed as vague and unreliable.” In one brought by the FTC the same year, the efficiencies that the court credited were not large enough on their own to decide the case in the merging firms’ favor.” The government lost both merger challenges but not as a result of the merging parties’ efficiency arguments.

REFERENCES


*United States v. Oracle Corp., 331 F. Supp. 2d 1998 (N.D. Cal. 2004); see Case 2 by McAfee, Sibley, and Williams in this part.


CASE 7


Serdar Dalkir and Frederick R. Warren-Boulton*

INTRODUCTION

On September 4, 1996, the two largest office superstore chains in the United States, Office Depot and Staples, announced their agreement to merge. Seven months later, the Federal Trade Commission voted 4 to 1 to oppose the merger on the grounds that it was likely to harm competition and lead to higher prices in "the market for the sale of consumable office supplies sold through office superstores." The merging parties chose to contest the FTC’s actions in court. On June 30, 1997, after a seven-day trial, Judge Thomas F. Hogan of the U.S. District Court for the District of Columbia agreed with the FTC and granted a preliminary injunction, effectively dooming the merger.

Staples broke new ground in terms of both the economic theory and the type of evidence presented at trial in an antitrust case. The antitrust enforcement agencies had traditionally focused on the increased probability of collusion following a merger as the primary theoretical underpinning for merger policy. In contrast, Staples spotlighted the potential for a merger to have "unilateral effects," a shift in focus first signaled by the 1992 revision

*Frederick R. Warren-Boulton served as an expert witness for the FTC in this case. Serdar Dalkir contributed to the economic analysis and the preparation of the expert testimony. Thanks are also due to Stephen Silberman, Robert Levinson, Melvin Orlans, James Fishlein, and Daniel Hosken for helpful comments on earlier drafts.
of the Department of Justice and FTC Merger Guidelines. Focusing on the characteristics of individual suppliers, the FTC argued that Staples, Office Depot, and OfficeMax were sufficiently different from other suppliers of office products, and sufficiently close competitors to each other, that the "sale of office supplies through office superstores" could be defined as a market separate from the sale of office supplies in general. In another departure, for evidence of the likely anticompetitive effect of the merger, the FTC relied primarily on direct estimates of the merger's effect on prices, rather than just predicting that an increase in seller concentration would cause significant (but vaguely specified) price increases. In addition to internal documents describing pricing policies and simple (but powerful) price comparisons between cities where Office Depot and Staples currently competed and those where they did not, the FTC's evidence on price effects included a large-scale econometric model that predicted the effect of the merger on prices. It also included an "event study" that used stock market data to calculate both the effect of the merger on shareholders and the financial market's implicit estimate of the effect of the merger on the prices charged by office superstores.

BACKGROUND

Office Depot and Staples are, respectively, the first- and the second-largest office superstore (OSS) chains in the United States. Staples pioneered the office superstore concept in 1986. In 1997, Staples operated approximately 550 stores in twenty-eight states. It had 1996 revenues of some $4 billion and a stock market valuation of approximately $3 billion at the end of 1996. Office Depot, which adopted the concept of superstores within months after Staples invented it, operated more than five hundred stores in thirty-eight states, had 1996 sales of approximately $6.1 billion, and had a stock market value of about $2.2 billion at the end of 1996. The rationale for the superstore concept was simple: While large businesses were able to purchase office supplies through high-volume contract stationers, small businesses and individuals had no comparably convenient, low-cost source of office supplies and other business-related products. The office superstore was to do for office supplies what the supermarket had done for home groceries.

The typical superstore is approximately 23,000 to 30,000 square feet in area, stocks 5000 to 6000 items, is located in an urban business area, and looks like a warehouse. Approximately half of Staples' and Office Depot's revenues are derived from sales of office supplies, with the rest coming from the sale of computers, office furniture, and other business-related goods.

items. Both chains purchase virtually all of their inventory directly from manufacturers in large quantities, enabling them to receive volume discounts that are unavailable to small and medium-sized retailers. These lower costs have led to dramatically lower prices: Office supplies are typically sold by superstores at discounts of 30 to 70 percent below manufacturer-suggested retail prices.

At one time, twenty-three competing OSS chains slugged it out in the market. By the time of the proposed merger, however, OfficeMax was the only remaining close rival to Staples and Office Depot. Spun off from K-Mart in 1994, OfficeMax operated 575 superstores and seventeen delivery centers in over 220 areas in forty-eight states. Like Staples and Office Depot, each OfficeMax superstore offered an extensive selection of over 7000 items at discount prices, selling primarily to small and medium-sized businesses, home office customers, and individuals. OfficeMax’s total revenues for fiscal year 1997 were $3.2 billion, with office supplies making up about 40 percent of total revenues.

The success of the OSS concept had redefined the retailing of office supplies in the United States, driving thousands of independent stationers out of business, just as the growth of supermarkets had driven out thousands of small “Mom and Pop” grocery stores. The competitive rivalry between the superstores had, however, benefited consumers substantially. Each OSS chain slashed prices; drove down costs; developed innovative approaches to marketing, distribution, and store layout; and expanded rapidly, bringing to increasing numbers of consumers the convenience of one-stop shopping at low prices. Office Depot had, at least in recent years, been the most aggressive and lowest-price competitor.

On September 4, 1996, Staples and Office Depot announced an agreement under which Staples would acquire Office Depot by exchanging 1.14 Staples shares for each outstanding Office Depot share, a roughly $4 billion deal. After a seven-month investigation, the FTC decided to challenge the merger.

THE FTC’S CASE

The FTC argued that this merger could be expected to lead to a significant decrease in competition in the market for consumable office supplies sold

After the Commission’s initial vote, the FTC staff negotiated a tentative agreement (subject to the Commission’s approval) with Staples and Office Depot that would have authorized the merger to proceed unchallenged if the two companies agreed to divest a sufficient number of stores to OfficeMax to preserve two competitors in cities where Office Depot and Staples were currently the only two superstores. On March 26, 1997, OfficeMax signed an agreement to buy sixty-three Staples and Office Depot stores for the fire-sale price of $108.75 million, subject to the consent of the FTC. But on April 4, 1997, the Commission voted to reject the proposed settlement and thus to challenge the merger.

Page 180
through office superstores and that the resulting price increases could be expected to be substantial. To prove its case, the FTC used a number of sources of data and analytical approaches to predict the price effects of the proposed merger. It argued that all of the evidence indicated that there would be large and long-lasting price increases, and therefore considerable harm to consumers.

The FTC was careful to compare the expected merger-related changes in prices and costs with the prices and costs that would have prevailed in the absence of the merger. Specifically, the FTC recognized that OSS prices might continue to fall after the merger but argued that because prices would fall significantly further without the merger, the merger would still harm competition. Likewise, the FTC stressed that the efficiencies claimed by the defendants must be merger specific.

**Concentration and the Competitive Effects of a Merger**

The underlying theme of merger policy is that mergers or acquisitions should not be permitted to create, enhance, or facilitate the exercise of market power, defined as the ability profitably to maintain prices above competitive levels for a significant period of time. The *Merger Guidelines* emphasize two ways in which mergers can lead to higher prices: coordinated interaction and unilateral effects.

When only a few firms account for most of the sales of a product, those firms can sometimes exercise market power by either explicitly or implicitly coordinating their actions. Coordinated interaction is of particular concern in homogeneous product markets, where all firms must charge very similar prices. Circumstances may also permit a single firm, not a monopolist, to exercise market power through unilateral or noncoordinated conduct, that is, without the concurrence of other firms in the market or in the absence of coordinated responses by those firms. Unilateral price effects are of particular concern if the products or services are differentiated, but those supplied by the merging firms are much closer substitutes for each other than for those of other suppliers. In any case, the exercise of market power causes a transfer of wealth from buyers to sellers and a misallocation of resources.

**Defining the Relevant Market:** "Consumable Office Supplies Sold Through Office Superstores"

The FTC argued that the relevant product market was "the sale of consumable office supplies through office superstores." The FTC supported its market definition, in part, by introducing evidence showing that: (1) OSSs offer a distinct set of products and services; (2) OSSs regard each other as their primary competitors; (3) non-OSS retailers do not tightly constrain OSS pricing; and (4) a hypothetical merger to monopoly among all three OSSs
could be expected to result in a significant increase in their prices for consumable office supplies—an outcome that would not occur if OSSs and other stores selling office supplies were in the same product market.

1. Office superstores offer a distinct set of products and services. The FTC argued that OSS firms were different from other vendors of office products because they carried a broad range of consumables and maintained large amounts of stock on hand. These attributes of office superstores created a one-stop-shopping opportunity for consumers that was not provided by other retailers or mail-order suppliers of office products.

Like customers of supermarkets and department stores, customers of office supply superstores benefit from being able to buy a large number and variety of products on a single visit. The full "price" to an office superstore customer of acquiring these products is the amount paid to the store, plus the customer's noncash costs of shopping. These noncash costs include the value of the time required to visit the store, gather information about products and prices, and shop. Since each visit to a store involves a fixed cost, customers prefer to purchase a bundle of items on each visit, especially low-cost "consumable" items that need to be purchased regularly.

Customers who purchase a bundle or basket of items need to decide: (1) which store to go to and (2) what products to buy on each visit. The first decision is relevant if one is analyzing a merger among a particular class of retailers (e.g., office superstores, department stores, or supermarkets) and needs to define a market for a particular type of retailing service. The second decision is relevant if one is analyzing a merger among manufacturers of particular products sold by those retailers (e.g., binders, women's dresses, or canned tuna).

OSSs devote significant shelf space to consumable office products and maintain a large inventory to ensure the convenience of one-stop shopping for customers. Superstores carry up to 7500 items of consumable office supplies, computers and computer-related products, and office furniture. While certain non-OSS retailers (mass merchandisers, warehouse club stores, computer stores, and consumer electronics outlets) sell a number of the same products that OSSs sell, they typically stock far fewer office supply items and/or carry a very limited assortment of consumable office supplies.

In court, both sides presented witnesses, exhibits, and affidavits that addressed the extent to which OSS retailers differ from non-OSS retailers of office supplies. Faced with a mass of conflicting evidence, the FTC strongly recommended that the judge visit several sellers of office supplies to see for himself how superstores differ from other office supply retailers. As one FTC expert witness put it, "One visit would be worth a thousand affidavits."

*Estimates of office supply items carried by the warehouse club stores range from 100 to 289. Mass merchandisers like K-Mart and Target typically carry fewer than 570 office supply items. Even Wal-Mart, which carries a relatively broad range of office supply items (between 1067 and 2400), nonetheless did not appear to be a significant competitor of the OSS firms.

Page 182
2. OSSs regard each other as their primary competitors. The parties' internal documents (at least those pre-dating the merger announcement) showed that each was concerned primarily or exclusively with competition from other office superstores. Indeed, Staples defined "competitive" and "noncompetitive" markets solely in terms of the presence or absence of OSS competitors, and referred to its participation in an "office superstore industry." Office Depot's documents similarly focused primarily on other OSS firms as competitors. The FTC argued that such evidence demonstrated that Staples and Office Depot recognized that other OSS firms were their main competitors.

3. Non-OSS retailers have little effect on OSSs' price changes. The FTC argued that the presence of non-OSS retailers could be expected to have little effect on the prices charged by OSS, especially in markets where more than one OSS was present. This implied that the presence of non-OSS retailers in an area would not prevent the merged office superstore from raising prices and that such non-OSS retailers should not thus be included in the relevant market.

The FTC did not dispute the fact that, in markets defined by Staples as "noncompetitive markets" (i.e., in markets where only one OSS was present), retailers like warehouse clubs and computer stores would be the closest competitors of the OSS. But the FTC argued that one could not infer from this that non-OSS retailers would provide effective competition for OSS firms in "competitive" markets, those where two or more OSSs already were present. A monopolist maintains a price so high that any further increase would cause a sufficient loss of customers to be unprofitable. Thus, a monopolist is distinguished not by the fact that it faces no competition, but by the fact that its closest competitors are too distant to prevent it from maintaining its price at a level significantly above cost. Ultimately, however, every monopolist "creates" its own "competitors" by maintaining its own price sufficiently high.'
Thus, in a market with two OSS firms, each OSS could overwhelm­
ingly be the other’s primary “competition” and provide the only effective
force holding the other OSSs’ pricing at present levels. If these two OSSs
merged, the new firm would find it profitable to raise its prices until com­
petition from non-OSS retailers eventually made further price increases
unprofitable. The postmerger OSS monopoly would then be constrained by
the prices charged by these new, non-OSS “competitors.” In short, even
though warehouse price clubs or Wal-Mart might be important competitors
to Staples in geographic markets that have no other OSS rivals, such non­
OSS suppliers are not significant competitors to Staples in geographic mar­
kets where Staples faces other OSS competitors, that is, in the markets that
the FTC thought were relevant to analyzing this merger.

The FTC’s econometric analysis supported the conclusion that non­
OSS competitors do not constrain OSS pricing in geographic markets
where two or three OSS chains are present. Indeed, simulations of the
effects of eliminating individual non-OSS retailers from such markets
showed that none of those retailers (except Best Buy, which had tried and
failed to implement an OSS-type format and had effectively exited by the
time of the merger) had any statistically significant effect on Staples’ prices.

Further evidence of differences between OSS firms and other office
supplies retailers involved price differences. In general, suppliers that com­
pete in the same market have similar prices for the same products. If con­
sumers can easily switch among suppliers, higher prices, adjusted for
quality, will not be sustainable. The FTC presented evidence that office
superstores in the same geographic market tend to price office products at
the same level, just as warehouse clubs in the same geographic market tend
to price office products at the same level. However, prices for office prod­
ucts in the same geographic market often differ significantly between OSS
firms as a group and warehouse clubs as a group.’

4. Econometric evidence supported an OSS product market. Under the
Merger Guidelines, the relevant product market in this case turned on the
following question: Would a merger to monopoly among the OSS chains in
a city allow the merged entity to raise the prices of consumable office sup­
plies by 5 percent or more? If the answer is yes, then “office supplies sold
through office superstores” is a relevant market under the Guidelines.

The FTC addressed this question by constructing a large-scale econo­
metric model of prices for office supplies. The analysis was designed to

7 When consumers are deciding among stores where they can purchase a group or bundle of products,
competing stores in the same market would be expected to show a very similar price index for a rep­
resentative basket of products, without necessarily showing very similar prices on individual items.

8 A Prudential Securities survey reported that in Detroit all three OSS firms had virtually identical
prices for the basket of office supplies sampled (total prices differed by from 0.4 percent to 2.8 per­
cent). In contrast, the price of a basket of items common to any of the three OSS firms and to Best
Buy was 18 percent to 19 percent higher at Best Buy (Prudential Securities, 1995, pp. 64, 67).
determine how Staples’ prices varied from one store to another as a function of the number of nearby Office Depot or OfficeMax stores, the number and identity of potential nonsuperstore rivals such as discount mass merchandisers or warehouse club stores, and differences in costs and demand conditions across local markets. The FTC had weekly data from the parties, for over eighteen months, covering more than four hundred Staples stores in more than forty cities. The data included prices for a large number of individual stock-keeping units (SKUs) as well as a price index for consumable office supplies.

The FTC’s analysis predicted that a merger to monopoly in markets where all three OSS firms were present would raise the price for office supplies sold through OSSs in those markets by 8.49 percent. Such an increase would not be possible if OSS firms were constrained by other retailers. These results confirmed that “consumable office supplies sold through office superstores” was a relevant market under the Guidelines criteria.

The Merger’s Likely Anticompetitive Consequences

The FTC argued that voluminous evidence—structural, documentary, and statistical—all supported the conclusion that the combined Staples/Office Depot entity would raise prices for office supplies. As to the structural evidence, a merger between the OSS firms in a hypothetical market with many OSS chains would not necessarily have any anticompetitive effect, because the merged firm would still have many close competitors. As we have seen, however, only three OSS chains compete anywhere in the United States. Therefore, OSS market concentration would increase significantly in all local markets in which both Staples and Office Depot were present as the number of OSS competitors fell from either three to two or from two to one. The companies’ own documents indicated that Office Depot was the main constraint on Staples’ prices and that, but for the merger, Staples planned to cut prices significantly over the next few years in response to current and future competitive pressures from Office Depot. The proposed merger would eliminate these pressures. Finally, statistical analyses of the potential effects of this transaction predicted that, absent efficiencies, the merger could be expected to lead to large price increases. In addition, data from financial markets indicated that investors implicitly believed the merger would lead to significantly higher prices even after allowing for the effects of any efficiencies.

The CEO of Staples, Tom Sternberg, testified to this point by arguing that, “Office Depot is our best competitor” and “our biggest competitor.” Sternberg described that this “best” and “biggest” competitor posed a more severe pricing constraint upon Staples than did the third office supply superstore chain, OfficeMax.

In fact, in anticipation of the merger, Staples canceled a 3 percent price cut on nonpaper supply items.
Structural Evidence: The Change in Concentration and Market Power

The structural effect of the proposed merger would have been to reduce from three to two the number of suppliers in markets where all three OSS firms would otherwise have competed and to create a monopoly in markets where only Staples and Office Depot currently competed, at least until entry by OfficeMax could reasonably be expected.

Table 7-1 shows Staples management's estimate for the percentage of Staples stores located in "Staples-only," "Staples and Office Depot," and "Staples, Office Depot, and OfficeMax" markets in 1995 and their projection for the year 2000. Absent the merger, Staples management anticipated a significant increase in competition from Office Depot and OfficeMax, as indicated by its projection that by 2000 markets with all three chains would account for 69 percent of Staples stores, up from 17 percent in 1995.

Therefore, the eventual effect of the merger would be to reduce the number of competitors from three to two in most geographic markets and from two to one in all but a few of the remaining geographic markets. (A small number of markets still would have only one OSS by 2000 even in the absence of the merger.)

Empirical Evidence Pointing to Likely Price Increases

In almost all merger cases before Staples, the DOJ or FTC relied primarily, if not exclusively, on indirect structural evidence of the kind presented above to infer that a significant price increase could be expected from that merger. Staples is unique, however, in terms of the large number of independent sources of strong, consistent, and direct evidence that were introduced at trial to show that prices would likely increase as a result of their merger. Five of these sources are discussed below.

**Predictions of Staples’ Management:** Staples’ own documents showed that, absent this merger, Staples’ management expected that wider competition would force it to lower prices and/or raise quality. Its 1996 *Strategy Update*, part of the FTC’s trial evidence, forecast that the percentage

<table>
<thead>
<tr>
<th>Year</th>
<th>Staples Only</th>
<th>Staples and Office Depot</th>
<th>Staples and OfficeMax</th>
<th>All Three</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1995</td>
<td>17%</td>
<td>29%</td>
<td>37%</td>
<td>17%</td>
<td>100%</td>
</tr>
<tr>
<td>2000</td>
<td>12%</td>
<td>7%</td>
<td>12%</td>
<td>69%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Source: Plaintiff's Exhibit 15, p. 32.
of three-player markets would increase to nearly 70 percent by the year 2000. It went on to predict that this could intensify the pressure on Staples’ prices and also lead to greater operating expenses as a result of a higher service quality and higher marketing expenditures.

Staples also predicted that, absent the merger, its retail margins, averaged over its entire sales (i.e., arranged not just over consumable office supplies and not just over markets where it faced competition from Office Depot) would decline by 150 basis points (”bps”), or 1.50 percentage points, by the year 2000 as a result of increased competitive pressure (ibid., p. 66). Of that margin fall, 60 bps would come from markets where Staples competed only with Office Depot and reflected Staples goal (absent the merger) to eliminate the price differences on nonpaper supply items between Staples and Office Depot.

Direct Comparisons of Prices across Local Markets: Statistical data generated during the ordinary course of business by the companies showed that, on average, both Staples and Office Depot priced significantly lower when they confronted each other in local markets.” As shown in Table 7-2, Staples’ office supplies prices were 11.6 percent lower in markets occupied by Staples and Office Depot than in Staples-only markets; they were 4.9 percent lower in markets with all three OSSs than in markets where Staples faced only OfficeMax. Competition between Staples and Office Depot also had a significant restraining effect on Office Depot prices. These data could be used to infer the likely increases in prices after the merger (on the assumption that Staples’ price patterns would dominate): +11.6 percent for the markets where premerger there was a Staples-Office Depot duopoly (accounting for 29 percent of Staples’ stores); and +4.9 percent for the markets where premerger all three OSSs were present (accounting for 17 percent of Staples’ stores).

Estimates from Econometric Analysis: The FTC performed an econometric analysis using store-level price data to estimate how prices differed across markets depending on the number and identity of firms in a market.” In essence, this econometric analysis was a more formal and

---

“In court, the FTC presented a particularly striking example of these price differentials: matching full-page color advertisements that appeared on the same day in two Florida cities, Orlando and Leesburg. Every detail was identical except the prices, which were 30 percent to 114 percent higher in Leesburg (with Office Depot only) than in Orlando (with three OSSs). This natural experiment provided the clearest evidence of both the existence of an OSS market and the likely effect of the merger on prices. To see a copy of the exhibit, go to [http://dalkir.tripod.com/depot/index.html](http://dalkir.tripod.com/depot/index.html). The statistical analysis was based on a large sample of store-level price data, drawn from 428 Staples stores in the United States over the twenty-three-month period from February 1995 to December 1996. The model examined statistically how Staples’ prices varied with the extent of OSS competition, the presence of non-OSS firms (such as Wal-Mart, K-Mart, Target, and Best Buy), and potentially location-specific cost and demographic variables. See Baker (1999), Ashenfelter et al. (2002), and Ashenfelter et al. (2006) for an extensive discussion of the econometric studies that examined the extent of localized competition between the merging firms. For a critique of the FTC’s pass-through estimates, see Werden, Froeb, and Tschantz (2005).
complete analysis of the kind of data just discussed. Using these estimates, the FTC calculated the overall price effects of the proposed merger: an average of 7.3 percent for the two- and three-firm markets where the merger partners were both present.

 Estimates from the Prudential Study: A Prudential Securities (1996) study reported the results of a pricing survey that compared prices for office supplies at office superstores in Totowa, New Jersey, a three-player market, and in Paramus, New Jersey, a nearby (twenty-five-minute drive) two-player market (Staples and OfficeMax). The survey showed that prices, especially on visible general office supply products, were more competitive in three-player markets than in two-player markets. In particular, the survey found that Staples’ prices on a basket of general office supplies that included the most visible items on which the office supply superstores typically offer attractive prices were 5.8 percent lower in three-player Totowa than in two-player Paramus.

 Estimates from a Stock-Market Event-Probability Study: Financial market investors vote with their dollars (or bet) on whether a merger will raise or lower prices. A merger that raises market prices will benefit both the merging parties and their rivals and thus raise the prices for all their shares. Conversely, suppose the financial community expects the efficiencies from a merger to be so large that the merged firm will drive down market prices. In this case, the share values of the merging firms’ rivals would fall when the probability of the merger goes up. Thus, evidence from financial markets can be used to predict market price effects when significant merger-related efficiencies are alleged.

The authors analyzed the effect of the proposed merger on share prices and concluded that, if consummated, the merger would raise the value of OfficeMax’s shares by 12 percent (or $200 million) but would have little or no effect on the share values of other retailers of office supplies; see Warren-Boulton and Dalkir (2001). These findings confirmed both that the

<table>
<thead>
<tr>
<th>Benchmark OSS Market Structure</th>
<th>Comparison OSS Market Structure</th>
<th>Price Reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Staples only</td>
<td>Staples + Office Depot</td>
<td>11.6%</td>
</tr>
<tr>
<td>Staples + OfficeMax</td>
<td>Staples + OfficeMax + Office Depot</td>
<td>4.9%</td>
</tr>
<tr>
<td>Office Depot only</td>
<td>Office Depot + Staples</td>
<td>8.6%</td>
</tr>
<tr>
<td>Office Depot + OfficeMax</td>
<td>Office Depot + OfficeMax + Staples</td>
<td>2.5%</td>
</tr>
</tbody>
</table>
merger would have anticompetitive price effects and that the OSSs constituted a relevant market."

Entry

Potential Entry of Other OSS Firms Did Not Constrain the Incumbents

The FTC argued that the threat of entry by a new OSS supplier would not prevent the merger from raising prices until such entry actually occurred. A potential entrant would assess the profitability of entry on the basis of what it expected prices to be after its entry, not before. Therefore, as long as incumbents could adjust their prices rapidly in response to entry, preentry prices would be irrelevant to the entry decision. And, since incumbents could not deter entry by keeping prices below the preentry profit-maximizing level, the best pricing strategy would be to "make hay while the sun shines." In other words, "investing" in entry deterrence by maintaining low prices was not a profitable strategy for incumbents.

Under certain conditions, however, potential competition can affect the prices of the incumbents. Usually, this requires both low sunk costs of entry and an inability on the part of incumbents to reduce their prices rapidly in response to entry. These conditions, however, were not present in the OSS industry. To the contrary, a significant share of entry costs into a local area was sunk costs, and incumbents could adjust their prices quite rapidly in response to entry. Therefore, prices of office superstore products could not be affected by potential entry.

This conclusion was supported by evidence in the documents. (For example, according to Thomas Sternberg, the CEO of Staples, Staples had not changed its prices in anticipation of entry by rivals.) The documents also showed that, when Staples considered entering a local market, it did not look at the prices in that market, but rather at the number of competitors.

Significant Barriers to Entry

While an individual office superstore could take advantage of store-level economies of scale and scope, a chain of superstores could also take advantage of economies of multistore operation. The latter economies appeared at different levels for different functions. Economies of scale in advertising, for example, clearly appeared at the local and regional levels.

"See Posner (2001, pp. 135-136) for a discussion of using event studies to establish indirectly the existence of efficiencies in merger cases.

If sunk costs are low (or firms are able to enter into long-term contracts with customers before actually entering) and incumbent suppliers are not able to change their prices quickly in response to entry, then the incumbents may not wish to encourage entry or to risk a significant loss of market share if entry occurs by maintaining high preentry prices.

Page 189
Thus, Staples’ strategy for entry into a large urban market consisted of first establishing a number of stores in the periphery and advertising only in local suburban papers until a critical mass was reached sufficient to make advertising in the large metropolitan newspaper or on television economical. For major markets, this implied a critical, minimum efficient scale of operation (a minimum number of stores) at the local level, with economies of scale for multistore operation that could extend into the regional level. The effect of such economies of scale on entry was described by Sternberg (1996, p. 59):

By building these networks [of stores] in these big markets like New York and Boston, we have kept competitors out for a very, very long period of time. Office Depot only came to metro New York in late 1995. They’re not in New York with any meaningful presence, they’re not in Boston, and they’re not in Philadelphia or anywhere in between. One of the reasons is that we have a very, very good network and it’s really tough to steal the customer from a direct competitor when you don’t have the economies of advertising leverage.

Sternberg’s description of Staples’ strategy to deter entry in its home base was similar: “Staples was trying to build a critical mass of stores in the Northeast to shut out competitors and make it cost-effective to advertise in the region’s high-cost media” (p. 61).

Some economies of scale in advertising even extend to the national level, perhaps due to a better ability to use network television advertising. Such economies give Staples a stronger incentive to enter markets where Office Depot and OfficeMax are already present, since this reduces advertising costs per dollar of revenues for Staples by increasing the total number of stores and the sales over which such costs can be spread.

All three OSS chains assess prospective new markets in terms of the existing numbers of OSS firms and the demand for additional OSS locations. Markets that have little or no “room” for additional stores are said to be “saturated.” Because multiple-store entry is typically necessary to enter a given metropolitan market, markets that are already saturated or nearly saturated are difficult to enter. An Office Depot document listed every market (as defined by Office Depot) in the United States and gave the total number of existing Office Depot, Staples, and OfficeMax stores, as well as estimates of the total number of OSS locations each area could support. The Office Depot estimates implied that, in many major markets in the United States today, there is insufficient demand for new office supply superstores to allow an entrant to achieve competitive-scale economies. In short, the time has fallen.

15 The parties defined “store potential” as the maximum number of OSS firms that can be supported in a given market, given existing market conditions, and defined the ratio of the number of OSS firms in a market to store potential as the degree of “market saturation.”
passed for a new chain to enter by building a significant number of stores in a new market without creating a glut of superstore capacity or locations. Thus, a firm currently attempting to enter cannot do so under the profitable conditions that the three incumbents faced in the past.

**Efficiencies Were Not Sufficient to Offset Price Increases**

The FTC argued that the efficiency claims made by the merger parties were exaggerated for several reasons. First, only efficiencies that are merger specific should be credited; that is, efficiencies likely to be achieved absent a proposed merger are irrelevant to the analysis of that merger. In this case, much of the anticipated efficiency gains were the result of the merged firm's increased scale. This in turn raised several questions: (1) Given the rate at which the parties were growing independently, many scale-related efficiencies could be expected in a short time through internal growth. (2) Achieving economies of scale in procurement does not require the expansion in retail operations that a merger would bring. Procurement cost reductions can be achieved by expanding sales through mail order or contract stationer operations, and both Office Depot and Staples had expanded such operations before their merger announcement. Thus, even if the parties had presented evidence to show that past expansions had lowered procurement costs, this would not have established that the claimed efficiencies were merger-specific. (3) Scale economies seldom continue indefinitely. Thus, particularly in the case of procurement costs, Staples and Office Depot may already be large enough to achieve the maximum sustainable price discounts that their suppliers can offer.

The second reason for the FTC’s skepticism as to the parties’ efficiency claims was the lack of support by reliable evidence. In particular, the efficiency claims made by the parties increased dramatically between the time that the deal was first approved by the Staples’ board and the time that the parties submitted an efficiencies analysis to the FTC. Because it was not clear what new information or insights the parties gained in that time period, there was a strong presumption that the substantially lower cost-saving estimates first presented to the Staples board were more reliable.

Third, under the Guidelines, efficiency gains are relevant only insofar as they result in a lower price to consumers. The share of any cost reductions that is passed on by a profit-maximizing firm increases with the proportion of those cost reductions that is attributable to variable (rather than fixed) costs; with the competitiveness of the industry; and with the share of firms in the market to which the cost reductions apply. In this case, the proposed merger would have substantially reduced competition. Further, any cost savings would have been limited to the merged firm. Therefore, historical estimates of the share of cost savings that the parties had passed on to consumers would significantly overstate the share of any merger-specific cost savings that would be passed on.
Specifically, the FTC’s analysis showed that the merger would bring true efficiencies that were the equivalent of only 1.4 percent of sales and that only one-seventh of these cost savings would be passed through to consumers. Thus, the net price effect of the merger would be substantial: the 7.3 percent price increase predicted by the FTC’s econometric model of pricing, less an efficiency pass-through of 0.2 percent (= 1.4% x 0.15), for a net increase of 7.1 percent.

THE DEFENDANTS’ ARGUMENTS

Staples and Office Depot argued that the merger would not have anticompetitive consequences. Their defense focused on two main arguments: (1) The FTC’s product market definition was erroneous; and (2) regardless of the market definition, the efficiencies from the merger, ease of entry into OSS retailing, and the defendants’ track record of lowering prices after their past acquisitions of other OSS firms all indicated that the merger would not raise prices.” Either of these two arguments, if accepted, would have disproved the FTC’s argument that the proposed merger would lead to a substantial lessening of competition in the relevant antitrust market.

Market Definition

The defense vigorously challenged the FTC’s claim that OSS firms constituted a relevant market for antitrust purposes. Staples and Office Depot argued that the FTC’s market definition was based exclusively on the identity of the seller and not on the characteristics of the product or service supplied by sellers. The respondents claimed that OSS firms were part of a broad market for retailing office supplies in which they held a low share. An OSS firm was constrained in its pricing not just by other OSS firms, but by all office product retailers.

The defendants argued that a retail product market is defined by the nature of the product being retailed; since office supplies sold by an OSS are not different from those sold by other retailers, both types of retailers are in the same market. The fact that OSS chains use different retail formats implies that they have found a particularly good way of competing with other retailers and does not imply that other retailers are in a different market. Thus, the defendants rejected the notion that office superstores supply a distinct bundle of goods and retail services that would enable a monopoly OSS to raise OSS prices.

“The defense cited two past acquisitions as examples of the two companies’ record of lowering their prices after a merger. According to the defense, the price of office supplies had fallen in each of the respective areas after Office Depot’s acquisition of Office Club in Dallas, Texas, and Staples’ acquisition of HQ Office Supplies Warehouse in Los Angeles, California (both in 1991).
The defendants also rejected the FTC’s argument that Staples’ and Office Depot’s own documents define OSS firms as "the competition" and "the market." Citing a previous court decision, they argued that the term market does not necessarily mean the same thing to a company and to an antitrust agency. Further, they contested the FTC’s use of selected passages in Staples and Office Depot documents as evidence in this regard: They claimed that other passages in the same documents used the term market also to include non-OSS firms. The defendants submitted exhibits showing that each regularly checked the prices of non-OSS firms, such as Wal-Mart, Viking, Best Buy, and CompUSA, along with the prices of other OSS firms. According to the defense, this illustrated the intense competition between OSS and non-OSS firms. As another illustration, the defense submitted a study that showed that the sales of a Staples store would fall by 1.4 percent with the opening of a new computer superstore, 2.4 percent with a new Wal-Mart, 3.7 percent with a new warehouse club, and 7.2 percent with a new Best Buy.

Efficiencies and the Net Price Effect

The defense claimed that OSS firms were founded on the principle of providing low prices through large sales volume. Thus, the defendants argued, the merger would increase the total volume of their (combined) purchases and lower the prices that they paid to manufacturers of office supplies. They also claimed that the merger would lower administrative, marketing, advertising, and distribution costs. Under the defense’s assumption that the merged entity would pass on to consumers two-thirds of the cost reductions, Staples and Office Depot would be able to cut prices significantly after the merger.

The defendants disputed the FTC’s argument that much of their claimed efficiencies could be achieved absent the merger. Moreover, they argued, even if some of those efficiencies could eventually be achieved through internal expansion, a merger would allow those efficiencies to be achieved much faster.

The defendants submitted an econometric study that suggested that Office Depot had a relatively small effect on Staples’ pricing and that a merger between the two would (absent efficiencies) increase prices for consumable office supplies by only 2.4 percent (compared with the FTC’s estimate of 7.3 percent) at Staples stores in markets with both Staples and Office Depot present, by 1.3 percent when averaged over office supplies at all Staples stores, or by 0.8 percent when averaged over all products and all Staples stores. The defendants also argued that, based on their estimate of cost savings and of the proportion that would be passed through to consumers (0.67 versus the FTC’s estimate of 0.15), the efficiency gains alone would cause prices to be lower by 3 percent over all Staples’ products and
stores. Thus, the net effect of the merger would be to reduce the prices faced by the average Staples customer by 2.2 percent (0.8 - 3.0% = -2.2%).

No Barriers to Entry and Ease of Expansion

The defendants argued that entry into the office supplies business was easy. Stores could be constructed within months, and sunk costs were low because the product did not decay and there were no fashion crazes. In addition, OfficeMax had increased its planned new store openings in 1997, demonstrating ease of expansion by existing competitors. Finally, entry or expansion did not necessarily entail costly new store openings: existing multiproduct retailers could enter, or expand into, the office supplies business by increasing the share of the shelf space they allocated to office supply items.

Public and Private Equities

The defense argued that blocking the merger would impose losses on both consumers and shareholders. The main consumer benefits from the merger that would be lost were the claimed efficiencies and lower prices discussed above; in addition, the combined company would be able to expand faster than either could individually, creating value for customers and for the U.S. economy. Any cost savings not passed on to consumers would benefit the shareholders of Staples and Office Depot. Finally, the defense argued that there was no need for a temporary restraining order or a preliminary injunction to stop the merger because the merger was reversible. If postmerger evidence demonstrated an anticompetitive effect, the merged entity could always be split back into two separate companies.

JUDGE HOGAN’S DECISION

The court agreed with the FTC and granted a preliminary injunction. Judge Thomas F. Hogan first noted that the law required the FTC to show only a reasonable probability of harm to competition to obtain a preliminary injunction. In his decision, Judge Hogan defined the relevant product market as the OSS submarket and found that Staples and Office Depot would have a "dominant market share" (between 45 percent and 100 percent) in

\[ \text{See Hausman and Leonard (1997); see also Werden, Froeh, and Tschantz (2005).} \]

\[ \text{Two examples offered to demonstrate the ease of entry were U.S. Office Products Co. and Corporate Express. Office Products had been founded recently (in 1994); both firms had expanded rapidly by acquiring small local dealers; their sales had also increased rapidly within the past few years.} \]

\[ \text{The defense's example was that Wal-Mart had already started expanding the shelf space it allocates to office products.} \]
many geographic markets after the merger. He also concluded that the FTC’s pricing evidence demonstrated a reasonable likelihood of anticompetitive effects.

The judge noted that neither the public nor the private equities claimed by the defendants were sufficient to offset the likely anticompetitive effects.

The Product Market

The court found that the sale of consumable office supplies by office superstores was a submarket within a larger market of all office supply retailers.” Baker (1997) discusses the judge’s opinion on the product market in light of the April 8, 1997, revised Merger Guidelines and concludes that the court’s “hidden opinion” treats the submarket argument as “a legal hook for reaching unilateral competitive effects from a merger among the sellers of close substitutes.”

Judge Hogan recognized that it was difficult to overcome the “initial gut reaction” to the definition of the product market as the sale of consumable office supplies through office superstores. Since the products sold by OSS firms are the same as the products sold by non-OSS retailers, “it is logical” to conclude that all these retailers compete. However, he noted, a firm could be a competitor in the “overall marketplace” without also being included in the relevant antitrust market.” He found plausible the FTC’s argument that a small but significant increase in one superstore’s prices would not cause a large number of its customers to switch to non-OSS retailers; instead, those customers would turn primarily to another OSS.

“The court stated that unscrambling the eggs, that is, undoing the merger if definitive anticompetitive effects were to be found in the future, was not a realistic option in this case. In addition to the difficulties involved in subsequently separating the merger partners, consumers would face the risk of being harmed if the merger was to be let through, and that damage could never be repaired by undoing the merger. Federal Trade Commission v. Staples, Inc., No. 97-701 (1997).

“In reference to submarkets within a market, the court decision cited the Supreme Court in Brown Shoe: Well-defined submarkets may exist that, in themselves, constitute product markets for antitrust purposes, and it is necessary to examine the effects of a merger in each such economically significant submarket to determine if there is a reasonable probability that the merger will substantially lessen competition. Brown Shoe defined several practical indicia to determine the presence of a submarket within a broader market, which Judge Hogan used to determine whether OSS chains constitute a submarket. See Brown Shoe v. United States, 370 U.S. 294 (1962).

“The court cited the notion of functional interchangeability in Du Pont (referring to interchangeability between cellophane and other wrapping materials) and Archer-Daniels-Midland (referring to interchangeability between sugar and corn syrup) cases. Noting that the Staples case is an example of perfect functional interchangeability in the sense that a legal pad sold by Staples or Office Depot is functionally interchangeable with a legal pad sold by Wal-Mart, it recognized that the analysis should go further and look at the cross-elasticity of demand between products, again citing the Du Pont case. See U.S. v. E. I. du Pont de Nemours and Co., 351 U.S., 377 (1956); and U.S. v. Archer-Daniels-Midland Co., 864 F.2d 242 (1988).

“The court did note some limitations of the data underlying the FTC’s individual analyses, and it further noted that the FTC could be criticized for looking at only brief snapshots in time or for
The judge observed that office superstores were very different from other office supply retailers in terms of appearance, size, format, the number and variety of items offered, and the type of customers targeted. While it was "difficult fully to articulate and explain all of the ways in which superstores are unique," he found that: "No one entering a Wal-Mart would mistake it for an office superstore. No one entering Staples or Office Depot would mistakenly think he or she was in Best Buy or CompUSA. You certainly know an office superstore when you see one." He argued that this is one practical indication for the OSS firms' constituting a submarket within a larger market.

Another practical indication for determining the presence of a submarket was "the industry or public recognition of the submarket as a separate economic entity." The judge found that the FTC had offered "abundant evidence" from the merging companies' internal documents that they evaluated their competition as other OSS firms and interacted with other OSS firms in making long-term plans. While Staples and Office Depot did not completely ignore non-OSS retailers, there was sufficient evidence that showed that Staples and Office Depot consider other OSS firms as their main competition.

**Likely Effect on Competition**

The judge was convinced that the proposed merger would likely have anticompetitive effects. He reached this conclusion from two pieces of evidence. First, having accepted the FTC’s product market definition, he found the concentration statistics to be a source of serious concern. After the merger, a combined Staples-Office Depot entity would have a dominant market share in many local geographic markets.

Second, the pricing evidence showed that an OSS was likely to raise its prices when it faced less competition from other OSS firms. Furthermore, without the merger, Staples and Office Depot would probably enter into each other’s markets and reduce prices. The merger would mean that these future benefits from increased competition would never be realized.

considering only a limited number of items, but it concluded that taken together, there was sufficient evidence for a low cross-elasticity of demand between the consumable office supplies sold by the superstores and those sold by other retailers.


"The premerger Herfindahl-Hirschman Index for the least concentrated market, Grand Rapids-Muskegon-Holland, Michigan, was close to 3600, whereas for the most concentrated market, Washington, D.C., the premerger HHI was about 7000."

"The combined market share would be 100 percent in fifteen metropolitan areas. In addition, in twenty-seven other metropolitan areas where the number of OSS firms would drop from three to two, the combined Staples-Office Depot market share would be above 45 percent. The HHI would rise on average by 2715 points because of the merger."
Entry

In a market defined as office supplies sold through superstores, the court focused on the entry of new OSS firms, not just any office products retailer. To achieve economies of scale and be profitable, a new OSS would have to open many stores and incur high sunk costs. Further, an entrant could not easily achieve economies of scale at the local level because many of the OSS markets were already saturated by existing OSS firms. The judge found it extremely unlikely that a new OSS would enter the market and counterbalance the anticompetitive effects of the merger.

Efficiencies

The judge noted that under the law it is unclear whether efficiencies constitute a viable defense. He stated that even if efficiencies can provide a legal defense in principle, in this case the defendants had not shown efficiencies sufficient to refute the FTC’s presumption of anticompetitive effects from the merger. He found that the defense’s estimates of the efficiencies were unreliable, unverified, and unrealistic. Among other problems, the defendants did not distinguish between merger-specific and other kinds of efficiencies, and given Staples’ historical pass-through rates it’s assumption that two-thirds of the cost savings would be passed through to the customers was unrealistic.

CONCLUSION AND AFTERMATH

The FTC’s victory in Staples came as a surprise to many observers. The casual empirical facts—there were many retailers of office supplies, and Staples and Office Depot together accounted for only a small percentage of the aggregate sales of such products—seemed determinative.

But the FTC’s careful marshalling of the data—especially, its use of the price data to show that the office superstores were a separate market—proved important in convincing the Commission itself and then Judge Hogan that the merger would be anticompetitive.

\[\text{As for the expansion of non-OSS suppliers into the OSS markets, the judge noted that it was unlikely that they would undo the merger’s anticompetitive effects. Specifically, the expansions by U.S. Office Products and Wal-Mart would be unlikely to constrain a potential increase in the prices of the merged entity. In relation to the defense’s argument that existing retailers could simply expand into the office products business by reallocating shelf space, the judge reasoned that while these retailers certainly had the power to do so, there was no evidence that they in fact would, following a 5 percent (small but significant) increase in the prices of the merged entity.}\]

\[\text{Historically, Staples passed through 15-17 percent of its cost savings to customers, as estimated by the FTC’s econometric analysis. For a discussion of the FTC’s estimation of the extent to which the merged firm would pass on cost savings from the acquisition to buyers, see Baker (1999).}\]

\[\text{White (2002) contrasts the court’s decision in Staples with the Surface Transportation Board’s decision in 1996 to allow the 3-to-2 merger of the Union Pacific and Southern Pacific railroads,}\]
Since Staples, both the agencies and merger applicants have routinely used direct evidence on the closeness of merging competitors and the expected size of a merger’s price effects, in defining the relevant product market and/or predicting the effect of the merger on consumers’ welfare."

In Staples, much of the efficiencies argument of the defendants was based on scale economies. Within three years following the merger’s abandonment, Staples and Office Depot each achieved the size (about 1000 stores) that they would have achieved as a single firm had the merger been approved (Balto 1999). As the parties’ premerger strategy documents had forecast, many of the new stores were in the overlap markets. As of March 2007, Office Depot had expanded to nearly 1200 office supply stores throughout the United States and Canada with worldwide sales of $15 billion. Staples had expanded to 1522 stores throughout the United States and Canada with worldwide sales of $16.1 billion at the end of its fiscal year 2005. OfficeMax continued as the third OSS, with 935 superstores in the United States (as of first quarter 2005) and worldwide sales of $13.3 billion in 2004. Thus, most of the efficiencies that the parties could have expected from the merger were achieved without much delay and without the detrimental price effects from a merger.

Although the growth of the Internet has resulted in new entry, as well as the creation or expansion of OSS firms’ online presence, the question as to what extent, if any, competition through the Internet has changed the OSS chains’ price setting behavior in local markets remains to be addressed empirically. And if “potential” competition (including the

with disastrous consequences for freight shipments in the U.S. Southwest in the late 1990s; see also Kwoka and White (2004).

"Scanner technology has dramatically increased the amount of data that is available for the analysis of mergers of companies with products that are sold at retail. One case that used these kind of data, FTC v. H.J. Heinz Co. and Milnot Holding Co. ("baby food case"), is the subject of Case 6 by Jonathan B. Baker in this Part. The FTC has also used scanner data in a similar way to evaluate supermarket mergers (see ‘‘A Blue Light Special for Mergers?’’ The Deal, October 5, 1999). The availability of scanner data has created a virtual ‘‘cottage industry’’ for econometricians predicting merger effects, either by estimating reduced-form equations as in Staples, or by a two-step process where demand elasticities are first estimated and then become inputs into a merger simulation model (see Werden 2002), an approach that has become so successful that it has created its own backlash (see Muris 2001)."

"Staples Annual Reports 2000 and 2005; ‘‘Staples Finds Office Market Staple: Chicago’’ (Chicago Sun Times, March 3, 2005); ‘‘Office Depot Acquisition Gains It Coffee Catch, Too’’ (Palm Beach Post, May 18, 2006)."

As an example, Sam’s Club (owned by Wal-Mart) launched a new online catalog for office supplies in 2005.

For example, Office Depot used its nationwide network of superstores, warehouses, and inventories to build the largest office-supply retailer on the Internet, including B2B (business to business) sales. Office Depot’s leveraging of its traditional network into the Internet guaranteed that its Internet business would be "profitable from virtually the day it opened" and successfully "neutralize[d] any potential threat from Internet startups" ("Why Office Depot Loves the Net: Its Brick-and-Mortar Network Is a Big Plus," Businessweek Online, September 27, 1999).

"Given the apparent strength of the economies of scope between Internet and brick-and-mortar retailing of office supplies, the Internet and traditional OSS operations may be more complementary than substitutable."

Page 198
Internet) has remained limited, a merged Staples-Office Depot would have continued to "make hay while the sun shines" in terms of higher prices in localized markets.

REFERENCES

PART II
Horizontal Practices
The Economic and Legal Context

Anticompetitive horizontal practices may occur in market settings ranging from fragmented industries to true monopolies. The nature of the practices, however, is likely to differ in each case. In an industry with many firms, the primary concern is with efforts at explicit collusion or implicit coordination in order to increase profitability above the competitive minimum. Where the number of firms is small—that is, in an oligopoly or dominant firm setting—cooperation or collusion remains a possibility, but anticompetitive conduct may include a richer array of strategies to disadvantage or drive out rivals or to deter possible entrants. Such strategies of predation and exclusion raise some of the most interesting and challenging antitrust issues.

The antitrust laws address this broad spectrum of anticompetitive conduct. Section 1 of the Sherman Act forbids any "contract, combination, . . . or conspiracy in restraint of trade . . ." language intended to prevent collusion among firms. Section 2 of that act prohibits actions that would "monopolize, or attempt to monopolize" a market. This is directed at acts designed unfairly to achieve or maintain market dominance. Section 5 of the Federal Trade Commission Act encompasses all of this in its sweeping ban on "unfair methods of competition."

Issues of horizontal practice differ from the problems addressed in Part I of this book. There the focus is upon market concentration and especially increases in concentration as a result of mergers that affect the intensity of competition in the industry. The practices discussed in this section are not the result of any structural change in an industry. Rather, they represent behavior patterns that arise within a given industrial structure as firms seek to increase their profitability either through closer cooperation with horizontal competitors or through aggression against one or more of them. We address these two categories of concern—cooperation and aggression—in that order.
CARTELS, COLLUSION, AND COOPERATION

A simple demonstration in microeconomics shows how a mutual agreement on price can increase profits, possibly up to the equivalent of the monopoly level. From a social perspective, this outcome has the same effect as shown in the introduction to Part I for the case of monopoly or merger with market power: Deadweight loss is created, and surplus is transferred from consumers to producers. Mutual agreement may mean a range of things—cartels, collusion, conspiracy and tacit cooperation. Cartels and collusion denote relatively formal agreements among sellers, while tacit cooperation involves informal coordination of behavior, which has uncertain legal implications. Conspiracy usually denotes something in between—more than a tacit understanding but less formal than a true cartel.

Formal cartels are prohibited in most countries except in limited circumstances such as export promotion, some agricultural marketing agreements, and certain sports league activities. Cartels among countries, however, are not subject to such prohibitions and have sprung up in many primary product markets ranging from coffee to crude oil. The experience of these cartels demonstrates that success is by no means assured; but when successful, they can be spectacularly profitable. The reasons for cartels’ varied successes bear examination, since the factors that prevent a cartel from succeeding must also stand in the way of an industry’s efforts at collusion or tacit cooperation. Put differently, if an industry cannot succeed in raising price through an explicit agreement, it is even less likely to succeed using less direct strategies such as tacit cooperation.

Broadly speaking, cartels must be able to (1) reach an agreement and (2) enforce adherence to the agreement. Economic theory and empirical evidence indicate that the following considerations with respect to the market, sellers and buyers, and demand conditions make efforts more likely to succeed in the effort to raise price and profitability:

1. the fewer are the number of firms;
2. the more similar are firm sizes;
3. the more difficult is entry;
4. the more similar are firms’ costs;
5. the greater is product homogeneity;
6. the smaller and more frequent are orders;


For a discussion of the rationale for these factors, see industrial organization texts such as Scherer and Ross (1990), Carlton and Perloff (2004), or Pepall, Richards, and Norman (2004).
• the more available is information about rivals' prices;
• the more stable is demand;
• the smaller and more numerous are buyers; and
• the lower are fixed costs.

The same factors play a role in firms' efforts to collude or conspire. As noted, however, conspiracy and collusion are considerably more challenging strategies for firms because of the difficulty of informal agreements and because of their illegality. Despite that, the appeal of higher profits prompts firms to enter into collusive agreements with surprising frequency.

The most common form of agreement among firms involves purely "tacit" modes of coordination. Much economic theory and empirical analysis have been devoted to examining the nature of firm interactions in general and the likelihood of spontaneous cooperation in particular. The relevant theory dates back to Augustin Cournot, who in 1838 demonstrated that above-competitive profitability can arise even in the absence of explicit agreements, so long as firms are sufficiently few and respond passively (and myopically) to each others' choices of quantities. Though Joseph Bertrand in 1883 demonstrated that in the case of price competition among firms selling identical products profitability would drop to competitive levels, even for a duopoly, more recent theoretical work has shown that even price rivalry among firms selling differentiated products will yield an equilibrium with above-competitive profitability.

While these theories generally predict above-competitive levels of price and profitability, they do not imply that monopoly levels are likely, since firms' individual incentives do not extend to that degree of cooperation. This proposition is readily shown with the aid of the famous game theoretic formulation known as the Prisoners' Dilemma. In the stylized example in Figure II-1, suppose there are two players A and B, each with price strategies we can label HI and LO. HI may usefully be thought of as the price consistent with cooperation or collusion, while LO maximizes each firm's individual profitability without regard to its rival or to joint profits. We assume that each firm makes its price choice independently but with knowledge that the resulting profit will also depend on its rival's price choice—the essence of oligopoly.

Player A rationally evaluates its two alternatives in light of what rival player B might do. If player B selects HI, then of A's two choices, its payoff
is greater from LO (15) than from HI (10). On the other hand, if player B selects LO, A is still better off choosing LO (4 versus 2). Remarkably, then, regardless of what A expects B to do, A earns greater profit by choosing LO! LO becomes what is known as A’s “dominant strategy”—better regardless of its rival’s choice of strategies. The same result holds for player B, so that the equilibrium of the game occurs at LO-LO despite the fact that each firm’s profit (and total industry profit) falls well short of its maximum at HI-HI. This outcome illustrates the powerful effect of competition—the risk that one’s rival will pursue its own interests—in conditioning the outcome of the process.

This compulsion toward strong competition is the result of the one-period (“static”) nature of the game. Most business interactions are not one-period, however, so game theory has also investigated equilibria in multi-period games, either of finite or infinite length. Treating a multi-period game as a sequence of one-period games with the structure shown in Figure II-1, it is straightforward to show that HI-HI may also be an equilibrium. To see this, consider player A’s choices beginning at that point. Player A could, of course, choose the LO strategy that was unbeatable in the one-period game, but if it did (and Player B chose HI), its rival would surely retaliate by also choosing LO in the following period. That would produce first-period profit for A equal to 15, followed by a never-ending series of 4, as shown in the time profile of profits in Figure II-2. The alternative is that player A sticks with HI, so that if B does as well, A earns a payoff of 10.

This holds for player B since the game described here is symmetric: Player B confronts exactly the same payoffs from its choices as does player A.

Infinite business games are not especially plausible either. Theory has shown, however, that finite games of unknown length are in fact much like those of infinite length. See, for example, Church and Ware (2000) or Pepall, Richards, and Norman (2004).

![Figure II-1: A “Prisoners’ Dilemma” Pricing Game](image)
indefinitely. Comparing the two time profiles, it is clear that unless player A’s discount rate is quite high—that is, unless it values first-period profits very much more than profits in later years—the longer period of modest profits from cooperation may well yield it a higher present value of profits. This result—that a cooperative equilibrium may result from purely tacit interactions—and the circumstances that may cause it are crucial ingredients in modern oligopoly theory.

Under some market conditions such spontaneous cooperation may not fully succeed, because of one distinctly unfavorable condition that prevents what otherwise would likely be successful cooperation. If that one condition can be remedied, higher profits may then result. Some methods of selectively remedying crucial conditions have come to be known as "facilitating practices" and are illustrated by each of the following:

- "Most favored customer" clauses, which reduce each seller’s incentive to cut prices to some customers by requiring equivalent rebates to other customers;
- Long-term customer contracts and exit fees from contracts, which may insulate existing sellers from new competitors and handicap existing firms in bidding away each other’s customers;
- Advance preannouncements of price changes or rapid ex post dissemination of such changes, which may assist oligopolists’ efforts to come to an agreement on and to maintain prices.

Each of these mechanisms may facilitate anticompetitive outcomes in the market. To the extent that they do, agreements sustained by facilitating

*In the present case cooperation is an equilibrium for any discount rate less than 20 percent. Circumstances characterized by high discount rates include companies facing imminent financial crises that need cash upfront, and countries with pressing development needs. Each of these cases is associated with players’ defecting from the agreement, as theory predicts.*

*A good analytical discussion of these various practices can be found in Salop (1986). See Hay (1999) for an illustration of facilitating practices in the Ethyl case.*
practices are in principle easier for policy to address than purely tacit agreements since the key facilitating practice can then be prohibited. Unfortunately, the dilemma for policy is that most facilitating practices also have benign or pro-competitive explanations. Disentangling the two remains a major challenge for economics and policy.

Three further considerations round out this brief review. First, collusion and coordination may occur not only with respect to price and quantity, but also with respect to capacity, product characteristics, advertising, research and development, and virtually any strategy variable. Most of the same broad considerations apply to these other dimensions of competition and cooperation, though details such as sunk costs, certainty, and time lags can play important roles (Scherer and Ross 1990). Second, perfect collusion or cooperation is not necessary for success. Sellers in an industry may be considerably better off simply by moderating the rivalry among themselves, even if full monopoly cooperation is not achieved. Third, higher price and profitability tend to spawn entry and erosion of market power. That outcome does not imply that the conspiracy was a failure, since the alternative may be zero profits, nor does it imply that policy has no role. The latter depends upon the relative timing of policy action versus the market-induced decline in market power.

Antitrust

The longest-standing precedent in all of U.S. antitrust policy has been the per se prohibition on explicit price fixing. Beginning with the Trans-Missouri Freight Association case and strongly affirmed in the Trenton Potteries decision, the Supreme Court has ruled that price fixing is an automatic violation of the Sherman Act’s ban on “contracts, combinations and conspiracies in restraint of trade.” The Court explicitly rejected the defense that the resulting price might be “reasonable.” This per se approach stems from a literal reading of Section 1 of the act, but it is also an efficient rule for certain types of conduct.

Per se prohibition is an appropriate policy stance in any of three circumstances: where the action always has adverse consequences; where it always has either adverse or at best neutral consequences; or it almost always has adverse or neutral consequences and its positive outcomes are both infrequent and difficult to distinguish. Each of the first two cases represents fairly clear grounds for a per se rule since neither prohibits beneficial actions. The third case is more problematic in that it concedes

9 U.S. v. Trans-Missouri Freight Association, 166 U.S. 290 (1897).
11 “The second may not be entirely free of controversy insofar as it occasionally attacks actions that are completely neutral in their market effects.
the possibility of occasional policy error—attacking beneficial actions. The justification is that these outcomes are relatively few and very costly to identify. A "perfect" rule in these circumstances would be very costly to administer (all possible cases would have to be fully examined), and the result might not be correct categorization in any event.

Given the earlier economic discussion, the Court’s per se approach to price fixing would seem appropriate. To be sure, the courts have muddied the waters a bit by some minor inconsistencies in their treatment of price fixing, notably in the Depression-era Appalachian Coals’ decision. More recently, however, the courts have indicated a receptivity to arguments as to the benefits of allowing direct competitors in certain circumstances to coordinate pricing and other matters. The specific circumstance is said to arise when coordination is required in order for a product simply to be offered at all. That is, transactions costs or some other impediment would otherwise cause a market failure. Beginning with the ASCAP/BMI cases, the Supreme Court has accepted this logical possibility and in those cases in fact allowed coordination among competitors.

One consequence of this exception to the per se ban is that companies being investigated for price coordination may now argue that they fall within the exception. In principle this might force the enforcement agencies and courts first to decide that issue before applying the per se rule in any matter, thus threatening to undermine the per se approach altogether. In an effort to prevent this administrative by costly outcome, the agencies have developed a so-called modified per se approach toward such practices. This approach asks a series of questions designed summarily to filter out all those practices that lack any such justification, treating them under the standard per se rule and leaving only a very small fraction of cases for a more comprehensive analysis.

If the effects of a particular practice depend crucially or entirely on the facts of each case, then a so-called rule of reason must be used instead of per se. The rule-of-reason is appropriate when a practice may have either adverse or beneficial effects, both in significant proportion, and when these alternative outcomes can be readily distinguished in actual practice. While information exchanges fall into this category, a brief review of the courts’
attempts to draw relevant distinctions illustrates the difficulties. For example, trade associations cannot take any action intended to force compliance by their members with published prices, but they are permitted to collect and disseminate considerable price information, some of it clearly helpful in efforts at coordination. Moreover, the courts have been somewhat more tolerant of exchanges of information about nonprice matters, apparently in the belief that these pose a lesser risk of coordination or that coordination is less likely to affect the market adversely.

Further difficultive have accompanied the courts’ recent efforts to judge so-called facilitating practices—those institutional arrangements and mechanisms that promote coordination among companies without explicit agreements. These cases have involved such practices as adherence to common pricing books that reduce the complexity of pricing heterogeneous products (as in the GE-Westinghouse case16), the use of most-favored-customer clauses as incentive-altering devices (as in the Ethyl case17), and reliance upon central databases to disseminate, virtually instantaneously, information about competitors’ prices (as in the Airline Tariff Publishing (ATP) dispute.18 The courts have been willing to consider some of these practices as undue intrusions on the market process, but their insistence on clear guidance regarding circumstances in which these represent anticompetitive behavior as opposed to normal business practice had been a common theme. They have often found such guidance lacking.

The courts have been extremely reluctant to convict companies for purely tacit collusion, since that comes close to prohibiting conduct that seems virtually inevitable in an oligopoly setting (for example, responding to a rival’s price change in a similar fashion). Moreover, prohibition is not straightforward since it is not altogether clear what action by companies should be proscribed. For these reasons purely tacit cooperation has been largely outside the reach of the antitrust laws and is likely to remain so unless in particular cases it is clear that some facilitating practice plays a crucial role and unless the courts are persuaded of the practicality of action against it.

**MONOPOLIZATION: PREDATION AND EXCLUSION**

**Economic Theory**

Oligopolists and dominant firms have a wider range of possible anticompetitive actions at their disposal than simply collusion and cooperation with

---

their rivals. In particular, they may pursue strategies designed to improve their own profitability by aggressively attacking their rivals (actual and potential) and thereby weakening their constraining influence. For example, when a firm willfully excludes potential entrants from its market, drives one or more of its present rivals out, or disciplines its rivals for their conduct, the extent of (long-run) competition in the market may be lessened. Such general statements, however, obscure the great difficulty in determining precisely which practices, and under what circumstances, should constitute antitrust violations. Indeed, the answers to these questions have changed over time and are still undergoing change as economic understanding advances. Our discussion here will focus on two categories of conduct that is associated with the "monopolization" offense in the law: predation and exclusion.

There is a longstanding and extensive body of antitrust law and economic analysis concerning predatory pricing. What might be termed "classic predation" involves a leading firm that increases output so as to lower market price and impose losses on its smaller rivals. Those rivals are eventually forced to exit, leaving the predator firm with unconstrained, or at least less constrained, market power. The new market equilibrium involves a yet higher market price and additional consumer harm. This scenario, together with other allegations of anticompetitive practices, has often been said to explain the rise of Standard Oil to dominance of petroleum markets in the late nineteenth century. There is no dispute that many of Standard's rivals exited the market and that it came to possess substantial market power, but a century after the fact, many continue to argue that Standard was simply a tough, efficient competitor and that its displacement of its smaller, less capable rivals was inevitable.

This debate typifies the controversy that has always surrounded allegations of predatory conduct. Chicago School adherents, in particular, have argued that the circumstances under which predation can succeed are very limited, that rational firms will therefore rarely attempt it, and hence that predation is not a significant policy problem. They note that the dominant firm's larger size implies that it will incur losses proportionally greater than those of the target firms, the proportion being its share relative to that of its rivals. Thus, the would-be predator is likely to injure itself more than it damages its rival. Moreover, even if successful in driving rivals out, the dominant firm will benefit from this strategy only if it has the protection of entry barriers in the postpredation period. Without such barriers, any effort by the firm to raise price will simply induce entry, and it will never recoup the losses that it initially incurred. These considerations, it has been argued, make predation rarely rational (McGee 1958).

This conclusion received support from some studies of the history of predating cases. Many instances of alleged predation have been found to
arise from rivals confronted by hard competition by more efficient larger firms, with adverse consequences for their own market share and profitability but not necessarily adverse consequences for competition. Some analysts therefore concluded that instances of true predation are sufficiently rare that they can be safely ignored—indeed, are best ignored, since efforts to distinguish predation from legitimate and beneficial price reductions would likely endanger the latter. While this inference goes too far for most observers, there is widespread agreement that true predation is considerably less common than alleged.

Modern economic analysis has advanced this debate by demonstrating that predation may be quite rational, and hence entirely plausible, under certain circumstances. The new analyses and based on models of information asymmetry and/or strategic behavior and fall into three broad categories: 1) First, a dominant firm may predote in one market or against one competitor in order to deter competitors in other markets in which it may operate, or in the future against competitors in the same market. Even if "irrational" when considered in isolation, such conduct may create a reputation for aggressive response that discourages any other competitors from initiating action. The value of that reputation justifies the expenditure in the initially targeted market.

Another possible mechanism arises if the dominant firm has a "deeper pocket"—that is, greater financial resources with which to battle its rival, due to differential access to capital markets. Differential access arises when small firms have to pay a premium to borrow funds, either because lenders favor the prospects of the leading firm or perhaps because the leading firm deliberately disrupts the business prospects of smaller rivals.

Finally, some economic models have shown that a dominant firm may use pricing in an effort to convince ("signal") actual and potential rivals that it has lower costs. The most straightforward way to send that signal is to pick a price consistent with being a low-cost firm—that is, a low price. Rivals will not know (or at least cannot be sure) that this signal is false, so they may decide that they would be better off tempering their actions or simply operating in some other market.

These new models of rational predation are central to any economic discussion of the topic, but none of them provides a simple and clear guideline for what constitutes predatory pricing. That is, practical policy is not easily based on such possibilities, although many have attempted to formulate decision rules to help guide policymakers (see, for example, Brodley et al. 2000). Most rules also allow for predation that varies from the classic model—for example, predation by a nondominant firm (Burnett 1999); or

"For an extended discussion, see Ordover and Saloner (1989).
predatory price cuts targeted at particular rivals (Pittman 1984); or predation involving strategic use of product innovation (Ordover and Willig 1981), product proliferation (Schmalensee 1979), product replacement (Menge 1962), advertising (Hilke and Nelson 1983), and cost manipulation (Salop and Scheffman 1983). Which of these might be employed depends on the circumstances—in particular, the aggressor firm’s perception of the most readily exploitable weaknesses of its rivals. All of these make clear how diverse predatory conduct may be, and how correspondingly difficult is the task of fashioning a comprehensive rule.

Closely related to the strategy of attacking present rivals are actions by a dominant firm or monopoly to exclude firms contemplating entry into a profitable market. There is a wide variety of such exclusionary practices, including: selective price discounting in niches or to customers most likely to be subject to entry; new product introduction and placement that make entry less attractive; long-term contracts with present customers to deny demand opportunities; tie-in sales or bundling; advertising directed at specific niches and potential entrants; denial of necessary inputs to rivals, and so forth. For these to be effective, economic theory teaches that the actions or threatened response must be credible; that is, the leading firm must actually follow through because it is rational for it to do so. Otherwise, the firm’s threats will not be taken seriously, and they will fail to deter entry. Credibility is enhanced by sunk costs, first move advantages, and other irreversible strategies.

The first of these models, and the ones that have continued to receive the most attention, are those examining how capacity expansion may rationally be used to preempt entry opportunities. Early work by Spence (1977) and Dixit (1980) showed how the incumbent may select capacity and creditly commit to a strategy that alters the best response of possible entrants. By leaving little opportunity for viable operation by an entrant, the result may be either no entry at all, or at worst entry on a smaller and less threatening scale. A host of such models has now been developed, examining various assumptions and parameters giving rise to credible entry deterrence and extending the results to the large number of other strategy variables noted above. The upshot of this research has been to place entry deterrence and related strategies on a sound theoretical footing and to provide some guidance as to the conditions that make them plausible.

Another potentially anticompetitive strategy is often said to involve "essential facilities," those for which access is necessary for effective

"Among relevant references, see Aghion and Bolton (1987) and Scherer (1979).

"For a good discussion of these, see Church and Ware (2000).

"For extended discussion, see Shapiro (1989) and Martin (1993)."

Page 212
operation in a market. Incumbents that deny prospective entrants such access can control the structure and conduct within the wider market dependent on that necessary input. Since essential facilities often (but not always) arise in the case of network industries, further discussion of the issues raised will be postponed until Part IV.

Quite a different category of conduct that raises possible concern involves exclusionary practices arising in related—often vertically related—markets. More will be said about these in Part III, but some issues deserve mention here as well. For example, a company whose product dominates one market may seek to tie or bundle the purchase of that product to another where it has no such dominance but is interested in establishing a presence. Such "leveraging" of market power from one market to another may displace rivals in the second, "competitive" market and raise barriers to subsequent entry. Although some economists had argued that leveraging cannot extend monopoly power, more recent models (e.g., Rasmussen et al. 1992) have demonstrated how leveraging may indeed have that consequence. The conditions under which this is a rational, profit-maximizing strategy for the incumbent firm, however, remain imperfectly specified in practice.

In sum, there is clearly an enormous variety of practices that dominant firms may employ to handicap both small rivals and entrants. Many of these may also have innocent purposes and raise no competitive concerns. It is an ongoing challenge for economics to determine the circumstances that distinguish pro- versus anticompetitive consequences.

Antitrust

Antitrust law has not been able to await resolution of these conceptual and practical difficulties before dealing with allegations of predatory and exclusionary behavior. Clearly, the difficulties of distinguishing predation from hard, but honest and pro-competitive, conduct by large firms can easily result in policy errors. Treating dominant firm actions too harshly is likely to stifle the very competition that antitrust policies were designed to foster. A policy that is too lenient, on the other hand, risks harm to competitors and consumers alike. At various times the enforcement agencies and courts may have made mistakes of both types.

Early cases against monopolies and dominant firms revealed the courts' inclination to interpret adverse effects on competitors as evidence of predation. In the previously mentioned *Standard Oil* case the dominant firm was alleged to have engaged in nasty practices toward rivals. Whether this was true or not, it was primarily on that basis that the Supreme Court found that Standard "monopolized" its market. But in cases where a firm's dominant position was achieved without victimized rivals, the opposite verdict
was reached. Thus, market dominance by U.S. Steel that resulted in higher prices was found not to violate the antitrust laws since there was no evidence of "brutalities and tyrannies" against smaller rivals. In fact—and not surprisingly—its higher prices drew favorable reviews from its competitors.

Judicial focus on and hostility toward the conduct of large firms may have reached their peak during the 1960s. Cases involving mergers, price discrimination, and monopolization routinely included allegations of predatory conduct by some leading firm in the market. In the *Utah Pie* case, for example, large food manufacturing firms were convicted of predatory price discrimination against a local firm despite the fact that the latter remained the market leader and was profitable throughout. Such decisions raised questions as to whether price competition itself was being sacrificed to protect smaller firms.

In an effort to advance the debate concerning predatory conduct, two leading antitrust scholars (Areeda and Turner 1975) published an analysis seeking to establish enforceable rules for predation. In their analysis, Areeda and Turner did not claim that predation would never occur. Rather, they sought to limit judicial prohibitions on low prices to those cases that they believed were demonstrably anticompetitive by contemporary economic standards. Their survey of various possible price and cost patterns led them to conclude that anticompetitive effects were likely only if the leading firm priced below its own marginal cost, and they proposed using average variable cost as a surrogate for marginal cost.

This Areeda-Turner rule now has many critics, but the criticism in no way detracts from the importance of its contribution. For one thing, in the confusion that had pervaded judicial opinions with respect to predation, the concreteness of the Areeda-Turner rule had an immediate attraction, and many (but not all) courts seized upon it for their own use. Moreover, that original rule caused an explosive growth in research attempting to clarify the economics of predation and to develop an operational rule to identify it. Unfortunately, the latter goal has proven elusive, and the major advances in economic understanding of rational predation have made virtually no impression upon the courts (Klevorick 1992). Instead, the Court has continued to insist on evidence regarding the relationship between cost and price, and more recently has added a virtual requirement that recoupment be shown possible."

Analogous difficulties have arisen with policy toward exclusionary behavior by dominant firms. Distinguishing legitimate conduct—even

---

"The language is taken from *Standard Oil*.


"See Elzinga (1999) on Matsushita."
hard competition—by a dominant firm from anticompetitive exclusion of potential competitors is as formidable a policy task as determining true predatory conduct. Examples of alleged exclusion include cases brought against such familiar dominant firms as Kodak, IBM, Kellogg/General Mills/General Foods, and DuPont. In Berkey v. Kodak, the plaintiff alleged that Kodak sought to monopolize film and photo finishing by its introduction of a new camera/film system without predisclosure to rivals in film and processing.

A whole series of cases brought by peripheral equipment manufacturers and later by the Justice Department against IBM alleged selective price discounting, strategic equipment modification, and other harmful practices designed to exclude potential entrants. In the breakfast cereal case the Federal Trade Commission (FTC) developed a novel economic argument that the three major manufacturers shared a monopoly and collectively "packed product space" with new cereals so as to leave inadequate opportunity for new entrants. The FTC also pursued monopolization allegations against DuPont for using its cost advantage in the market for the paint whitener titanium dioxide to preclude expansion by rivals.

With minor exceptions all of the defendant firms in these cases prevailed. The court ruled that Kodak had no obligation to predisclose product innovations for the benefit of rivals. IBM won most of the myriad private suits that it faced, and the Justice Department ultimately withdrew its suit. The FTC acquitted the three breakfast cereal manufacturers, in part balking at the proposition that "more products" could somehow be anticompetitive. In the DuPont case, the FTC concluded very explicitly that a dominant firm is under no obligation to avoid hard competition against smaller or newer rivals, regardless of the fate that might befall these rivals.

The view that a dominant firm should have much the same degree of discretion in its behavior as a competitive firm is not well-founded in economics. The same behavior may be benign when employed by a small firm but have quite different consequences in the hands of a dominant firm. Policy shifts in this direction relaxed restraints on the conduct of dominant firms, and the number of such cases—both government and private—dwindled sharply during the 1980s. More recently, however, the enforcement agencies and courts have on occasion been persuaded of instances in which dominant firms' actions have crossed the line.

26 Berkey Photo v. Eastman Kodak Co., 603 F.2d 263 (2d Cir. 1979).
27 See Brock's (1989) discussion of several cases brought by peripheral equipment manufacturers against IBM.
28 In re Kellogg et al., FTC Docket No. 8883 (1981).
29 For an extended analysis of this case, see Dobson et al. (1994).
This appears to be the case in both the Kodak and Microsoft proceedings discussed in Part IV of this volume. These and other cases may imply greater enforcement and judicial receptivity to claims of anticompetitive actions based on the focused application of economic theory and evidence.

REFERENCES


---

30 See Case 21 by Jeffrey MacKie-Mason and John Metzler in Part IV. This is a different case against Kodak than the one previously discussed.

31 See Case 20 by Daniel Rubinfeld in Part IV.


CASE 8
Predatory Pricing in the Airline Industry:

*Spirit Airlines v. Northwest Airlines (2005)*

Kenneth G. Elzinga and
David E. Mills

INTRODUCTION

Spirit Airlines is a regional airline carrying passengers to and from a small number of cities in the United States. Northwest Airlines is a major airline carrying millions of passengers each year all around the world. Almost everyone reading this chapter has heard of Northwest Airlines; most readers will never have heard of Spirit Airlines. Such is the disparity in size and reputation between the two firms.

In 1996, Spirit and Northwest tangled in a price war on two domestic routes that each served: Detroit-Philadelphia and Detroit-Boston. Fares dropped to levels that most travelers between these cities never thought possible. Spirit claimed that Northwest’s fares were so low as to be predatory; that these prices would lead to Spirit’s exit from the markets; and that when that happened, Northwest would raise fares to monopoly levels and consumers would be harmed. Northwest responded that the low fares in these two markets reflected head-to-head competition between the airlines. According to Northwest, consumers were the beneficiaries of the low fares.

In 2000, Spirit filed an antitrust suit against Northwest, alleging that Northwest’s tactics were predatory and violated Section 2 of the Sherman Antitrust Act. Northwest challenged this claim, arguing in court that

*The authors were consultants to Spirit Airlines. They are grateful to Federico Ciliberto and the editors for comments.*
Northwest’s prices met neither an economic nor a legal definition of predation. Northwest added that to use the antitrust laws to penalize the kind of price competition that took place between Spirit and Northwest would harm consumers and set a dangerous legal precedent for thwarting vigorous price competition in the future. A federal district court agreed with Northwest that Spirit’s claims lacked merit; the U.S. Court of Appeals for the Sixth Circuit disagreed and sent the case back to the district court for a trial.

The case of Spirit Airlines v. Northwest Airlines provides a classic opportunity not only to examine the economic character of price competition between these two firms but also to study the nature of competition between airlines. Even beyond that, Spirit Airlines v. Northwest Airlines offers a peg on which to hang the entire issue of predatory pricing. Thus, we start with the law and economics of predatory pricing. Then we turn to the antitrust battle between Spirit and Northwest.

PREDATORY PRICING

Predatory pricing is a business strategy designed to create or maintain a monopoly position. In the canonical predatory pricing episode, a monopolist uses temporarily low prices to drive out a small incumbent or thwart an entrant whose presence would bring competition into the market. Here, in theory, is how it works: The predator cuts its prices to impose financial stress upon a small rival or an entrant and to induce the target firm to exit the market. Once the target withdraws, the predator can charge supracompetitive prices.

In the short run, during the price war, consumers enjoy low prices. But consumers end up paying monopoly prices in the long run. If the predator’s discounted long-run monopoly profit exceeds its short-run profit sacrifice during the price war, then consumers’ discounted long-run loss will exceed the gain that consumers enjoy during the price war. All things considered, consumers would be better off and the market would be more efficient if the target firm survived rather than being eliminated in a price war.

Some theories of predatory pricing incorporate reputation effects and other strategies that exploit informational asymmetries by the predator. For example, a firm that aggressively cuts prices in response to entry may gain a reputation for being a “bully,” which could deter other entrants in the future or in other markets. When part of a predator’s payoff from an episode of low prices comes from establishing or strengthening a reputation for responding aggressively to future entry, the firm’s payoff in the instant market need not offset its short-run profit sacrifice solely in that market.¹

¹Bolton et al. (2000) suggest that strategic considerations like reputation effects occur more frequently than courts have recognized. For a different perspective, see Elzinga and Mills (2001).
Predatory pricing presents a policy dilemma for antitrust: The conduct of a predator during a predatory pricing episode looks a lot like vigorous competition, so it is hard to distinguish an instance when low prices are sustainable and pro-competitive from an instance when low prices are unsustainable and seek to persuade the target firm that its prospects for commercial success are poor. Ordover and Willig (1981) identify predatory pricing as an episode of low prices that would not be profitable for the monopolist unless they cause the target firm to exit the market (or, in some instances, accede to the monopolist’s price leadership).

The dilemma for antitrust in identifying predatory pricing is heightened because the positive payoff from constructive intervention (in a predatory episode) and the negative payoff from obstructive interference (in a pro-competitive episode) both are greater early in the episode, when telling the difference is the hardest. The best time to arrest a predator is early on, before the entrant leaves the market. But compelling evidence of predation is scarce early on, and the risk of spoiling competitive price-cutting is greater when antitrust puts a quick end to low prices instead of waiting for the entrant to fail.

Two decades ago, in his influential Supreme Court opinion in Matsushita, Justice Lewis F. Powell observed that "there is a consensus among commentators that predatory pricing schemes are rarely tried, and even more rarely successful." This view was informed by the theoretical contributions of Bork (1978) and McGee (1958, 1980), among others, who concluded that, as a rule, predatory pricing is irrational. Subsequent theories of rational predatory pricing that incorporated various informational asymmetries (including, for example, the reputation motive noted above) between a predator and preyed-upon firms appeared in the economics literature. These are summarized by Ordover and Saloner (1989) and Klevorick (1993). While this literature questions the view that predatory pricing is uncommon, little has happened in the courts in the years since Matsushita that would challenge Justice Powell’s generalization—and, indeed, it was reinforced by the Supreme Court in 1993 in its Brooke Group decision and more recently in its 2007 Weyerhaeuser decision.

Some commentators on price wars in the airline industry have suggested that this industry may be an exception to Justice Powell’s “rarely tried” observation. For instance, Alfred E. Kahn (1998, p. 1), who played a prominent role in the deregulation of the U.S. domestic airline industry, stated that “[t]here are in fact strong reasons to believe that, at least so far as

---

4Weyerhaeuser Co. v. Ross-Simmons Hardwood Lumber Co., Inc., 127 S. Ct. 1069 (2007). This was a case involving allegations of predatory bidding.
the airlines business is concerned, the Supreme Court’s view . . . is simply incorrect.” Characteristics of the airline industry that may predispose it to predatory price wars include the multiple city-pair markets in which airlines face each other, the price discrimination strategies available in the sale of tickets, and the unusual cost structure of airlines.

Whether and how frequently incumbent carriers respond to entrants in a predatory fashion is a question that recurs regularly in the antitrust community. Bamberger and Carlton (2006) provide an account of how incumbent airlines have responded to the entry of so-called low-fare carriers since 1990. Although their review “finds no evidence that predation by major carriers against low-fare entrants . . . was systematic and successful during the mid-1990s,” when the growth of low-fare carriers slowed temporarily, they “cannot rule out isolated instances of predatory conduct . . .” (2006, p. 22).

During the 1990s, as the low-fare carrier segment of the U.S. airline industry expanded, several entrants complained to the U.S. Department of Transportation (DOT) about allegedly predatory conduct by incumbents.’ The DOT investigated these claims and released a report agreeing that incumbent airlines had responded to low-fare entrants with predatory practices and proposed specific guidelines for delineating unlawful conduct (USDOT 1998). These proposals prompted Congress to commission a study by the Transportation Research Board of the National Research Council, which several months later issued a report (TRB 1999) that reiterated the DOT’s concerns, albeit with some reservations.

In May 1999, the Department of Justice (DOJ) sued American Airlines for predatory conduct against low-fare carrier entrants at American’s Dallas-Fort Worth hub. A district court dismissed the case in 2001 on several grounds. The court rejected the DOJ’s various applications of the “average variable cost test,” as well as the DOJ’s recoupment theory, which was based on the alleged predator’s acquiring a reputation for predation. The court also reasoned that since American’s low fares only matched, and did not undercut, the low-fare entrants’ fares, American was entitled to “meet” its new competition. The Tenth Circuit Court of Appeals affirmed the trial court’s decision in 2003.’

A significant part of the DOJ’s failure to win American was because of the court’s reasoning as to how a price-cost test and a recoupment test should be administered in a predatory pricing case in the airline industry. Notwithstanding the legal outcome, American did not persuade everyone that if an incumbent airline (1) cut its fares to match a low-fare entrant,

Bamberger and Carlton (2006, p. 2) report that the DOT received thirty-two such complaints between March 1993 and May 1999.


United States v. AMR Corp, 335 F. 3d 1109 (10th Cir. 2003).
(2) increased its capacity sharply to reduce the number of occupied seats on
an entrant's flights, and (3) then sharply raised fares and reduced capacity
once the entrant retired, then this conduct should be outside the reach of the
opportunity for a trial court to distinguish predatory conduct from legiti­
mate competition in this industry. Before turning to the facts and an analy­sis of the *Spirit* case, we review the legal standard that currently applies to
predatory pricing claims.

**DISTINGUISHING PREDATORY PRICING
FROM COMPETITION**

The most recent predatory pricing opinion issued by the Supreme Court
involved two cigarette manufacturers. Neither firm was a new entrant to
the cigarette industry. Brown & Williamson's aggressive pricing of its
newly introduced discount cigarettes prompted Liggett, who previously
was the main supplier of discount cigarettes, to cry foul. In a lengthy
antitrust battle that made its way to the Supreme Court, the Court, relying
on economic analysis, set a high bar for plaintiffs to succeed in predation
cases. According to the Supreme Court's *Brooke Group* opinion, proving
a predatory pricing episode requires three analytical steps.

The first is to define the relevant market in which the alleged predation
takes place and determine whether structural conditions in that market
make it plausible that, but for the target firm, the defendant firm could
exercise monopoly power. If the defendant's rivals in the relevant market
are sufficiently small and limited in their ability to discipline the defen­
dant's prices, and if there are significant barriers to entry in the market, the
necessary monopoly power is established. The analysis then moves to
the second step. But if monopoly power in the relevant market is not
established, economics teaches that it is implausible for the defendant's
aggressive pricing to harm competition in that market regardless of the
particulars. Consequently, the possession of monopoly power in the rele­
vant market, or a reasonable prospect of acquiring it, is a threshold issue in
proving that a defendant's pricing is predatory.

The second step in the economic analysis of predation is to determine
whether the defendant's prices in the relevant market were below an appro­
priate measure of its costs. The most prominent cost benchmark for decid­
ing predatory pricing cases is the Areeda-Turner (1975) test. This test asks
whether the defendant's prices were below the firm's reasonably anticipated

* * *  

case is often called *Brooke Group* because The Brooke Group acquired Liggett in the course of the
litigation. See also *Burnett* (1999).

Page 223
average variable costs (AVC). If the defendant’s prices were below this benchmark, Areeda and Turner suggested that the defendant’s low prices could be anticompetitive since such prices were not remunerative in the short run and were below the firm’s shut-down point. In addition, prices this low would be below the shut-down point of the defendant’s alleged target if that firm were no more efficient than the defendant. These prices might drive a rival out of the market and pave the way for a period of unimpeded monopoly pricing. Areeda and Turner proposed that charging prices above AVC would be a safe harbor for defendant firms. If the second step in the inquiry indicates that the defendant’s prices were above the cost benchmark, then the defendant’s pricing is presumed to be nonpredatory. Unless additional factors warrant continuing the analysis, the predatory pricing hypothesis is rejected. If the defendant’s prices were below the benchmark, predation becomes a more likely hypothesis. The analysis proceeds to the third and final step.

The third step investigates the likelihood that the defendant could charge monopoly prices high enough and long enough after disposing of the target rival to recoup the losses it endured during the period of low prices. This step is known as the “recoupment test.” Although showing that a defendant’s prices were below an appropriate measure of its costs usually is necessary, it is not a sufficient condition for proving predatory pricing. The recoupment test creates a safeguard for courts against making Type I errors (i.e., reaching false positive conclusions) in applying a cost-based test. Errors of this kind could occur if the defendant’s costs were measured improperly or if the defendant’s losses were inadvertent instead of strategic. The likelihood that the defendant’s low prices were predatory is sharply reduced if the recoupment test indicates that the payoff from ousting a rival would be insufficient to warrant the losses attending the low prices. Under the antitrust laws, to prove that the defendant’s low prices were predatory, a plaintiff must show that there was a “dangerous probability” that the defendant could recoup its predatory losses by charging monopoly prices after the rival withdraws from the market.

The plaintiff’s allegations in Spirit Airlines v. Northwest Airlines are similar to those in Brooke Group in that there was no reliance on predatory theories premised on the exploitation of information asymmetry between the firms. Given this similarity, we shall use the Court’s three-step Brooke

9 Katz (2006, p. 6) suggests that the recoupment test is a “reality check: If there was no reason for the firm that has been alleged to have engaged in predation to expect to be able to recoup, then it raises the question of why the firm would have ever tried to engage in predatory pricing.”

10 In addition, there may be limited circumstances that would induce a defendant to charge prices temporarily below benchmark costs for reasons that are not anticompetitive. For example, a firm may use low introductory prices while introducing consumers to a new product. Or the firm’s production learning curve may be so steep that it charges prices below cost for a time while costs are falling with accumulated experience.

BRIEF CHRONOLOGY OF EVENTS

Northwest Airlines was founded in 1926 as an air mail carrier between Minneapolis and Chicago. The firm began passenger service soon thereafter and today is one of the five largest passenger airlines in the world. As the firm’s operations at Minneapolis grew, Northwest developed a passenger service hub there. In 1986, Northwest merged with Republic Airlines, which had hubs at Detroit, Memphis, and Minneapolis. Northwest’s domestic hub-and-spoke route structure remained centered on these three airports during the period relevant to this case. Northwest also offers service to many international destinations both directly and through its global airline alliances with European and Asian carriers. In 2000, Northwest had about 53,500 employees, $11.4 billion in revenues, and net income of $256 million. That year the carrier reported approximately 58.7 million flown passengers.

Spirit Airlines is a low-fare carrier whose economic profile is very different from Northwest’s. In 2000, Spirit operated twenty-eight aircraft and had almost two thousand employees serving 2.8 million flown passengers. At the time of the events that are the focus of this case, Spirit offered regularly scheduled air service to thirteen city pairs. Its primary business was flying passengers from Detroit and other points in the Northeast to Florida and other entertainment destinations.

Prior to December 1995, Northwest’s share of the nonstop local passenger airline service between Detroit and Philadelphia (DTW-PHL) was about 70 percent. US Airways accounted for the rest of the local service on the DTW-PHL route. Northwest’s average (one-way) fares for local passengers on the DTW-PHL route for every month during 1995-1997 are shown in Figure 8-1.


"A local passenger on the DTW-PHL route is a passenger whose original point of departure is either Detroit or Philadelphia and whose final destination is the opposite city. Passengers, for instance, who fly from Minneapolis to Philadelphia via Detroit, or who fly from Philadelphia to Columbus by way of Detroit, are connecting passengers on the DTW-PHL route.

"Source: Data Bank 1A Superset. Carriers other than Northwest and US Airways accounted for less than 4 percent of the DTW-PHL market.

"Northwest’s average fares are constructed from Northwest’s Local Passenger Data for the DTW-PHL route. Unless otherwise noted, all fares reported in this chapter are one-way fares.

Page 225
In December 1995, Spirit introduced a single daily nonstop flight on the DTW-PHL route using an eighty-seven-seat jet aircraft. Spirit's average fares on this route, also shown in Figure 8-1, were substantially lower than Northwest's. Northwest's lowest unrestricted fare on the DTW-PHL route prior to Spirit's entry was $355, and its lowest restricted fare was $125. Spirit offered fares as low as $49 when it began service and, three months later, was selling tickets for an average fare of $57.

Initially, Northwest did not materially reduce its fares in response to Spirit's entry. However, when Spirit introduced a second daily nonstop flight on June 28, 1996, "Northwest immediately and dramatically reduced its fares, offering a lowest published fare that matched Spirit's $49 one-way fare, and . . . expanded its capacity on this route by adding another flight." As a result, Northwest's average fare fell to $77 in July 1996, as shown in Figure 8-1. Northwest's DTW-PHL fares were "less than Spirit's lowest fare for 92.5 percent of the days during the [alleged] predation period."

As Northwest increased its capacity on the DTW-PHL route, the number of passengers whom Spirit served dropped, and its operations on the route became unprofitable. Spirit achieved load factors as high as 88 percent on this route before Northwest responded to its entry, but load factors fell to

*Northwest defines three types of local passengers: true local, connecting local, and multidestination local. True local passengers are nonstop local passengers.

Source: Northwest Airlines, Spirit Airlines.

**Spirit Airlines' Response to U.S. DOJ Antitrust Division Civil Investigative Demand No. 17702, p. 5.

"DOT Presentation Outline, Northwest Airlines, Inc., p. 29.

"Spirit Airlines' Response to U.S. DOJ Antitrust Division Civil Investigative Demand No. 17702, Attachment 2, p. 12.


between 31 and 43 percent in subsequent months. Spirit canceled one of its DTW-PHL flights on August 20, 1996, and canceled the other on September 30, 1996 (Kahan 1998, p. 5). After Spirit canceled its flights, Northwest raised its fares sharply, as shown in Figure 8-1, and reduced capacity.

A similar sequence of events occurred on a second route. Prior to April 1996, Northwest was the only airline offering nonstop local passenger airline service between Detroit and Boston (DTW-BOS). Northwest's only competition on this route consisted of flights by other carriers that required a connection between DTW and BOS. Figure 8-2 shows Northwest's average fares for local passengers on the DTW-BOS route for every month during 1995-1997. Northwest was charging $411 for an unrestricted ticket on this route, and its lowest restricted fare was $129.

On April 15, 1996, Spirit introduced a single daily nonstop flight using an eighty-seven-seat jet aircraft on the DTW-BOS route. Spirit's introductory fare was $69, and its average fare for as long as it operated on this route was in the range of $67-75, as shown in Figure 8-2 (Kahan 1998, p. 7). Northwest responded to Spirit's entry by adding two additional daily nonstop flights on the DTW-BOS route and offering a matching fare. One of these flights used a larger aircraft than Northwest had used on the route previously. Northwest's sharp fare reductions are shown in Figure 8-2.

*Northwest defines three types of local passengers: true local, connecting local, and middestination local. True local passengers are nonstop local passengers.

Source: Northwest Airlines, Spirit Airlines.
These fares were "less than Spirit's lowest fare for 93.9 percent of the days during the [alleged] predation period."

After Northwest took these actions, Spirit’s load factors on the DTW-BOS route were between 17 and 31 percent, and its operations on the route became unprofitable. The company canceled its flight on September 8, 1996 (Kahan 1998, p. 8). Shortly after Spirit canceled its DTW-BOS flight, Northwest raised its fares on the route, as shown in Figure 8-2, and reduced capacity.

ECONOMIC ANALYSIS OF NORTHWEST'S CONDUCT

Seeking to establish that Northwest’s conduct in these two city pairs was predatory and violated Section 2 of the Sherman Antitrust Act, Spirit’s economic analysis followed the three-pronged approach suggested by the Court’s Brooke Group opinion and described earlier. Northwest’s economic analysis sought to discredit Spirit’s claims and prove that Northwest’s conduct was not predatory. In keeping with Ordover and Willig’s (1981) definition of predation, Northwest’s analysis was framed as a test of whether Northwest’s response to Spirit’s entry on the two routes was motivated by profit considerations that were attributable to injuring competition. The implementation of this test involved steps that were similar to those undertaken by Spirit. However, because Northwest’s analysis concluded that the carrier’s revenues on the routes in question were not below any relevant cost benchmark, Northwest claimed there was no need to examine whether Northwest had monopoly power or the ability to recoup predatory losses in the relevant markets.

Market Definition

The first step in Spirit’s analysis was to define the relevant markets and assess their structure. This was done to determine whether Northwest possessed the monopoly power necessary to execute a successful predatory campaign. Market share matters because a firm with a small share of the market would not have the pricing discretion required to execute a predatory pricing campaign all by itself. In antitrust economics, relevant markets have two dimensions. One identifies the relevant set of products that is bought and sold in the market. The other identifies the relevant geographic extent of sales.

Spirit took the position that the relevant geographic markets in this case were for passenger air travel between the city pairs of DTW-BOS

27Ibid.
and DTW-PHL. Northwest agreed with this position. Defining geographic markets as city pairs is plausible since an increase in the ticket price of a flight between Detroit and New York would have very little effect (if any) on the number of passengers who want to fly between Detroit and Boston. Defining geographic markets as "city pairs" assumes that the substitutability of travel between city pairs is de minimis compared with the substitutability of travel within city pairs.

Defining geographic markets as city pairs is generally accepted in airline industry circles. For example, Morrison (1998, pp. 150,156) states that "the unit of output of a passenger airline is transportation of passengers between cities" and it "... is at the route level, after all, that airlines actually compete with one another." The Transportation Research Board of the National Research Council (1999, pp. 65-66) states that "[a]irlines compete for passengers at the city-pair level. There are thousands of combinations of origin and destination (O-D) points that comprise the markets for air transportation ..." In comments before the U.S. Senate, a former head of the Antitrust Division (Klein 2000, p. 21) stated, ". . . relevant airline markets are likely to consist of scheduled airline service between a point of origin and a point of destination, generally referred to as city pairs. This market makes intuitive, as well as economic, sense."

While Northwest agreed with Spirit that the relevant geographic markets in the case were the DTW-PHL and DTW-BOS city pairs, Northwest took sharp exception to Spirit's proposed relevant product markets. This difference of opinion about product markets had much to do with the companies' disparate conclusions about Spirit's allegations.

Spirit offered two alternative definitions of the relevant product markets in this case. One alternative defined the markets to consist of all local passengers flying on the DTW-PHL route and, similarly, all local passengers flying on the DTW-BOS route. The other alternative defined relevant product markets comprising only leisure-oriented price-sensitive local passengers on the two routes. Spirit claimed that price-sensitive passengers are an identifiable class of travelers strongly prone to alter their travel plans in response to changes in fares. These are vacation travelers and travelers flying for personal reasons who are more flexible about scheduling and who pay for tickets out of their own pockets. According to Spirit, business-oriented non-price-sensitive passengers, who prize flexibility in their travel plans and are on expense accounts, generally purchase higher-price tickets with no restrictions. Price-sensitive passengers often purchase lower-price tickets with restrictions (e.g., advanced purchase, not fully refundable, or requiring a Saturday layover).

Both of Spirit's product market alternatives limited the relevant markets to local passengers on the DTW-BOS and DTW-PHL routes; passengers whose trips begin or end elsewhere were excluded even though they occupy seats on the same plane as local passengers. Purchasing a DTW-PHL ticket is a poor substitute for a passenger whose trip begins in
Indianapolis and ends in Philadelphia even if this trip involves a stopover in Detroit. Also, given the distances involved, Spirit’s product market alternatives limited the relevant markets to airline travel only and excluded other modes of travel between the cities in question. Although some passengers might drive to their destinations, take a bus, travel by train, or even charter a plane, Spirit deemed these alternatives as not being “reasonably interchangeable” with scheduled airline service.

Northwest disagreed with Spirit’s position that the relevant product markets included only local passengers and contended that connecting passengers on the DTW-BOS and DTW-PHL routes also should be included. The fungibility of seats on the same plane occupied by passengers with different points of origin and or destination, Northwest claimed, meant that product markets should include all passengers who traveled these routes. Northwest’s reasoning on this point stressed the carrier’s yield management system and supply-side substitutability more than demand-side substitutability. The companies also disagreed that the distinction between price-sensitive and non-price-sensitive passengers identified separate product markets.

Spirit cited several factors in support of defining the product markets to include all local passengers. Available data make this approach more convenient since isolating price-sensitive passengers from the rest requires drawing a line between passengers on the same plane traveling from and to the same cities. Defining all local passenger product markets also acknowledges supply-side substitution between passenger classes. An airline might choose to change, on short notice and at very little cost, the mix of seat prices on a given plane serving a particular route. For these reasons and perhaps others, the conventional approach taken in the economics literature does not distinguish price-sensitive and non-price-sensitive product markets.” In recognition of this convention, Spirit used an all-local-passenger market definition to assess Northwest’s monopoly power and pricing conduct. Spirit contended, however, that defining the product market expansively to include all local passengers exaggerates the substitutability of alternative tickets for many passengers who happen to be flying on the same route. One indication that many passengers in the two geographic markets do not regard all flight options as substitutes is the bimodal fare structure that Northwest maintained in the months immediately prior to Spirit’s entry. As shown in Figure 8-3, two-thirds of Northwest’s passengers in the DTW-BOS market during the preentry period bought tickets in a lower ticket price range. The remaining one-third were in a much higher price bracket. Figure 8-4 indicates that the same general pattern holds for Northwest’s fare structure in the DTW-PHL market.

“See, for instance, Borenstein (1989). However, Borenstein does exclude connecting or “through” passengers from the routes he investigates.
Spirit claimed that Northwest's bimodal fare structure in these markets illustrates weak substitutability between ticket options and supports its position that there are separate product markets for price-sensitive and price-insensitive passengers. If all passengers were in the same product market, Spirit reasoned, one would expect the price differential to be

FIGURE 8-3 DTW-BOS Northwest’s Fare Distributions for Single-Coupon Local Passengers Period before the Alleged Predation 1995Q2-1995Q4

Note: Data were not available for 1996Q1.
Source: DBIA.

FIGURE 8-4 DTW-PHL Northwest’s Fare Distributions for Single-Coupon Local Passengers Period before the Alleged Predation 1995Q3-1995Q2

Note: Data were not available for 1996Q1.
Source: DBIA.
compressed or arbitrated away. The fact that the price differential is so large suggests that the same flight serves at least two product markets.

The distinction between more and less price-sensitive passengers is recognized by the airline industry and is often couched in terms of "business" versus "leisure" travelers. For example, a U.S. General Accounting Office study (2001, p. 17) found that the structure of fares in the airline industry depicts markets for differentiated products and noted that "[p]assengers on the same commercial airline flight—sometimes even those in adjoining seats—may pay fares that vary widely. . . . The industry generally segments potential passengers into two categories—those traveling for business purposes and those traveling for leisure purposes." In its own analysis in a prior antitrust case, Northwest characterized differences in ticket prices on the same flight as stemming from differing elasticities of demand associated with "business" versus "leisure" travelers. As Northwest's executive vice president-marketing (Levine 1993, p. 25) once testified in another matter, "The airline industry generally believes that business (basically nondiscretionary) travelers constituted a somewhat distinct market. . . ."

Northwest disagreed that there were separate product markets for price-sensitive and non-price-sensitive passengers. In reference to the wide variation in the carrier's preentry fares, Northwest observed that after Spirit began service on the two routes, Northwest's fare structure became more compressed so that the carrier could retain passengers. Also, noting that Spirit's fares on these routes were unrestricted, as were Northwest's higher fares (i.e., those fares that Spirit claimed non-price-sensitive passengers paid), Northwest argued that Spirit competed just as much for passengers who previously paid higher Northwest fares as it did for those who previously paid lower fares. Spirit recognized that a significant number of its passengers on these routes might have paid higher Northwest fares if Spirit's service were not an option.

Monopoly Power

In their antitrust treatise, Areeda and Hovenkamp (2002, p. 341, footnote omitted) state, "... predation is usually assumed to be possible, if at all, only for a dominant firm with a very substantial market share. Given the variety of factors affecting the minimum size for plausible predation, no single market share number can be offered as a guaranteed minimum." Spirit argued that Northwest's shares of the DTW-BOS and DTW-PHL markets were substantial enough to support its claim that Northwest possessed monopoly power in those markets, however product markets are defined. In the year 1995, prior to Spirit's entry, Northwest had about 90 percent of the

"Areeda and Hovenkamp (2002, p. 341) hazard "[nevertheless, we suggest a strong presumption that Sherman Act predation is unreasonable at market shares below 60 percent."
DTW-BOS market and about 69 percent of the DTW-PHL market, as measured by the number of nonstop or one-stop local passengers flown. US Airways accounted for the rest of the local flights on these routes. In 2000, after Spirit had entered and exited those markets, Northwest had a market share of about 86 percent in DTW-BOS and about 64 percent in DTW-PHL. Whether calculated by assets, revenues, seat-miles, or number of employees, Northwest’s operations out of its Detroit hub dwarfed Spirit’s (and that of all carriers) during these years.

A large market share is a necessary but not a sufficient condition for the monopoly power required to be a successful predator. Entry conditions are important as well because predatory pricing is unlikely to be successful in a market where entry or reentry is relatively easy. Spirit identified several barriers to entry in the DTW-PHL and DTW-BOS markets and argued that these barriers supported its claim that Northwest had monopoly power in those markets. Some of these barriers stemmed from Northwest’s size and incumbency. For instance, Northwest enjoyed a degree of “brand loyalty” from many consumers because of the company’s frequent flyer programs and promotional programs for business travelers and travel agents. Spirit or any other entrant into the relevant markets in this case would have to overcome these incumbency advantages in order to compete successfully with Northwest.

Even though the mobility of planes and crews from one route to another generally facilitates entry in the airline industry, a limited supply of ground-based facilities can be an economic barrier to entry even if planes and crews are readily available. Important examples include ticket counter space, boarding gates, and takeoff and landing slots. Spirit claimed that these limitations supported its claim that Northwest had monopoly power in the relevant markets.

According to a U.S. General Accounting Office (1996, p. 9) report, “In 1990, our survey of the 66 largest U.S. airports revealed that 85 percent of their gates were leased to established airlines under long-term exclusive-use leases. At some airports, every gate was under an exclusive-use lease. We concluded that such leases limited entry because, in order to gain access to the airport, a nonincumbent would generally have to sublease gates from the incumbent airlines—often at less preferable times and at a higher cost."

“Source: Data Bank 1A Superset. Other carriers than Northwest and US Airways accounted for less than 4 percent of the DTW-PHL market.


“The scarcity of gates and take-off and landing slots is generally recognized as a barrier to entry in passenger airline markets. See, e.g., Morrison (1998, p. 166 ff)."
than the incumbent pays on the master lease.\footnote{See also Transportation Research Board (1999, p. 88 ff).} The Transportation Research Board (1999, p. 120) observed that “[w]hen few or no extra gates are reserved by the airport operator for other airlines, recurrent gate scarcity can become an impediment to entry. To prevent a new rival from gaining a foothold, it is possible that an incumbent airline might withhold gates or offer unattractive terms.”

The GAO identified DTW as one of the most gate-constrained airports in the United States. Of the eighty-six jet gates at this airport in 1990, seventy-six were under exclusive-use leases. And sixty-four of the gates at DTW were “leased to Northwest until the end of 2008, with all but 10 under exclusive-use terms” (USGAO 1996, p. 10). The vice chairman and general counsel of Spirit Airlines (Kahan 1998, p. 9) testified before the Transportation Subcommittee of the Senate Appropriations Committee in 1998 that “Spirit Airlines emplaned over 19,000 passengers from Detroit in December 1997 without a gate. We go from one carrier to the next seeking unused space for which we may contract at odd times of the day. Because we lack a gate … under the rules of the local airport authority … we are assessed a 25 percent surcharge over the rates charged to other carriers including, obviously, [Northwest]. The scarcity of gates at DTW, and Northwest’s dominance of existing gates, raised a barrier to entry into either of the relevant markets.”

In sum, Spirit argued that structural conditions in the relevant markets—Northwest’s market shares and entry conditions—indicated that Northwest possessed monopoly power in those markets sufficient to warrant advancing the economic analysis of Northwest’s conduct to the second step.

Northwest did not challenge Spirit’s market share calculations and did not undertake a detailed structural analysis of the relevant product markets. But the carrier disputed that it possessed monopoly power in those markets. It cited US Airways as a competitor sufficiently formidable to prevent Northwest’s exercising monopoly power in either market, especially DTW-PHL.

Northwest also contested Spirit’s claims that entry barriers such as airport facilities were sufficient to impede the entry of rivals in either market. It argued that Spirit had demonstrated its ability to contract for airport facilities at DTW, PHL, and BOS and was in a position to expand service on DTW-PHL or DTW-BOS by substituting the airport facilities that Spirit used for serving other routes. Northwest noted that Southwest Airlines began operating at DTW in 1987, although it did not serve either of the relevant markets in this case. Also, in 2000 Pro Air started operations at a secondary Detroit airport and offered service to PHL until it was shut down by an order from the Federal Aviation Administration. Noting that neither PHL nor BOS was considered to be a gate-constrained airport, Northwest reasoned
that the Pro Air experience indicated that Spirit could have expanded its service to those airports by using the secondary Detroit airport instead of, or in addition to, DTW (USGAO 1996, p. 9).

Analysis of Northwest’s Revenues and Costs

The second step in a predatory pricing investigation involves assessing whether Northwest’s fares in the DTW-PHL and DTW-BOS markets were below the firm’s average variable costs. Finding that Northwest charged fares below this benchmark would significantly raise the likelihood that its conduct in response to Spirit’s entry was predatory. In its regular course of business, Northwest produces a monthly record of the company’s financial and operational activity on specific flights and aggregates these data by city pairs.” Both Spirit and Northwest used these data in their analyses of Northwest’s revenues and costs.

Spirit’s Analysis of Revenues and Costs

A firm’s variable costs, as distinguished from fixed costs, are those costs that vary as the firm’s output level varies. Put differently, and in context, the relevant variable costs in a predatory pricing investigation are those that the alleged predator might avoid by curtailing production of that quantity of output associated with its challenged activity.” More descriptive terms than “variable” for these costs are “avoidable costs” or “incremental costs.” To assess Northwest’s conduct in either of the relevant markets alleged by Spirit requires comparing the fares that Northwest charged the relevant passengers with the avoidable or incremental costs associated with Northwest’s response to Spirit’s entry.

Northwest classifies its flight-specific costs into three categories: passenger variable expenses, nonpassenger variable expenses, and fixed expenses. Passenger variable expenses are costs that vary with the number of passengers on a given flight. These include, for example, costs associated with issuing tickets, processing passengers through the gate, in-flight food and beverages, and insurance and other liability expenses. Nonpassenger variable expenses are those that Northwest incurred when adding an additional flight or substituting a larger plane on an existing flight. These include items like the cost of fuel, pilots, maintenance, and landing fees. While these items cannot be adjusted as quickly or incrementally as passenger variable costs, they are not invariant to observed changes in passenger

“The source of these data was Northwest’s Flight Profitability System (FPS), which is “the primary means by which Northwest evaluates its financial performance...” “Northwest Airlines, Inc.’s Motion for Summary Judgment,” July 5, 2002, p. 11.

volume on a given route. Fixed expenses, like corporate overhead, are the costs of Northwest’s operations that are unlikely to change when additional passengers and flights are added to the carrier’s service on a given route.

Although Northwest aggregates its flight-specific costs across all passengers (e.g., connecting and local), Spirit reasoned that Northwest’s per-passenger costs were the correct measures to use for local passengers since there was not a material difference in the cost to Northwest of flying connecting versus local passengers on a given route. Northwest contested this assumption, noting that all flight-specific costs other than passenger variable costs are joint and common costs that should not necessarily be attributed pro rata to all passengers.

Spirit’s incremental cost calculations included all of the variable expense items identified by Northwest’s cost data. They also included certain items that Northwest classified as “fixed” costs. A critical item here turned out to be the cost of aircraft. Pacific could add seats to its local service on a route in three ways: increase the number of local service seats on existing flights; switch to larger aircraft; and add flights and aircraft. Since Northwest used the same types of aircraft on other routes as it used on DTW-PHL and DTW-BOS, the company could and did consider alternative uses of those aircraft when it deployed them on flights in the DTW-PHL and DTW-BOS markets. This meant, in the short run, that the cost to Northwest of offering service in one market was not offering service in another market. Accordingly, Spirit reasoned that the costs of aircraft used in Northwest’s challenged activity were incremental in the markets where this activity occurred. That is, even though these costs may be fixed insofar as the airline as a corporate entity was concerned, Northwest could avoid incurring these costs in the DTW-PHL and DTW-BOS markets by using them in other markets instead.

Spirit used a similar line of reasoning to estimate and allocate the rental value of airport gates, counter space, and similar airport assets to the incremental cost category. All of these assets were fungible and could readily be switched to other flights that originated or ended in DTW, PHL, or BOS. Northwest’s response to Spirit’s entry in DTW-PHL and DTW-BOS involved significantly increasing its passenger capacity on those routes. Spirit reasoned that the opportunity cost of aircraft and other airport assets that were diverted from uses on other routes in order to increase capacity on DTW-PHL and DTW-BOS were incremental to Northwest’s challenged activity. Northwest, by contrast, contended that these cost items were not incremental but instead were fixed because they were avoidable only in a longer time frame than genuinely variable costs.

*For a discussion of how to treat aircraft and other capacity-related costs in predatory pricing analyses in the airline industry, see Kaplan (2000). Kaplan was a consultant to Spirit in this antitrust case.
TABLE 8-1
Comparison of Northwest's Revenues and Costs ($/passenger)

<table>
<thead>
<tr>
<th></th>
<th>DTW-PHL</th>
<th>DTW-BOS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>July-September 1996</td>
<td>April-September 1996</td>
</tr>
<tr>
<td><strong>Incremental Cost include Aircraft</strong></td>
<td>53.47-60.17</td>
<td>65.87-85.24</td>
</tr>
<tr>
<td><strong>All Local Passengers Markets</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incremental Revenue with $49 Fare</td>
<td>44.29</td>
<td>n.a.</td>
</tr>
<tr>
<td>Incremental Revenue with $69 Fare</td>
<td>58.31</td>
<td>61.98</td>
</tr>
<tr>
<td><strong>Price-Sensitive Local Passengers Markets</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incremental Revenue</td>
<td>50.35</td>
<td>64.82</td>
</tr>
</tbody>
</table>


*This excludes incremental facilities costs.

The period of time after Northwest expanded capacity and cut its fares in response to Spirit's entry in DTW-PHL was July to September 1996; in DTW-BOS it was April to September 1996. Spirit's estimates of Northwest's incremental cost of passenger service in the DTW-PHL and DTW-BOS markets are reported in Table 8-1. In DTW-PHL, these costs ranged from $53.47 to $60.17 per passenger during the months of July, August, and September 1996. These costs included Northwest-designated variable expenses plus aircraft expense. If other airport expenses related to increasing capacity on the route were included, Northwest's incremental costs would be higher. The corresponding monthly incremental costs in DTW-BOS during April to September 1996 ranged from $65.87 to $85.24. These costs represented a substantial fraction of the company's average total costs in those markets during this time period. In previous litigation against American Airlines, Northwest had claimed that variable costs represented more than 95 percent of total costs.

As shown in Table 8-1, Spirit compared Northwest's incremental costs in the DTW-PHL and DTW-BOS all-local-passenger markets with the carrier's incremental revenues in those markets. Spirit argued that the appropriate fares to consider when calculating Northwest's incremental revenues in the all-local-passenger markets were the fares that were commonly paid by those Northwest passengers drawn to the carrier when it

"The fact that a predator does not react immediately to a new entrant or other competitive threat suggests that predation is a costly strategy. A rational predator will want to weigh carefully the expected consequences before incurring the costs of expanded output and below-cost pricing.


^Ibid.

reduced fares and added capacity. In DTW-PHL, these fares were $49 in July and part of August 1996 and $69 in the remainder of August and September 1996. In DTW-BOS, these fares were $69 from April through September 1996. These are not the fares that every Northwest passenger paid to fly on these routes, but most did: approximately 70 percent of DTW-PHL passengers and 75 percent of DTW-BOS passengers. Spirit reasoned that this was a sufficiently large share of the local traffic to justify assuming that Northwest’s incremental capacity was added to accommodate these passengers as part of Northwest’s alleged predation strategy.

Northwest’s incremental revenues were less than these fares by the amount that the company paid travel agents to sell their tickets. Netting out these commissions (and making other small adjustments for minor sources of revenue linked to the number of passengers flown), Northwest’s incremental revenues are shown in Table 8-1. The net revenue gained from the sale of a $49 ticket in the DTW-PHL market was $44.29. For a $69 ticket in the DTW-PHL market, it was $58.31. The net revenue gained from the sale of a $69 ticket in the DTW-BOS market was $61.98.

Comparing Northwest’s incremental revenues and costs in the DTW-PHL market shows that incremental revenues fell short of incremental costs during the July to mid-August 1996 period when Northwest offered the $49 ticket. Also, the incremental revenues from Northwest’s $69 ticket during the mid-August through September 1996 period were less than the company’s incremental costs of $60.17 in September 1996. In DTW-BOS, Northwest’s incremental revenues fell short of the company’s incremental costs in every month from April to August 1996.

Spirit made a similar revenue-cost comparison for the price-sensitive local passenger markets. Because Northwest’s financial and operational data do not distinguish price-sensitive and non-price-insensitive passengers (or even business versus leisure travelers), Spirit had to devise a method to identify price-sensitive passengers on the two routes so that their fares could be calculated. The conspicuously bimodal fare distributions in the two markets shown in Figures 8-3 and 8-4 enabled a plausible allocation of passengers to the price-sensitive and non-price-sensitive passenger categories for the preentry period in both markets. But during the postentry months when Northwest was allegedly charging predatory fares, Northwest’s bimodal fare structure collapsed, and many non-price-sensitive passengers took advantage of widely available low Northwest fares.

To identify the fares paid by price-sensitive passengers in the two markets in the alleged predatory period, Spirit sought to identify non-price-sensitive passengers and remove the fares that they paid from the rest. Spirit identified the number of non-price-sensitive passengers in the preentry period based on Northwest’s bimodal fare structure in that period. Then it extrapolated that number to the alleged predatory period, making an adjustment because demand elasticity would draw more non-price-sensitive passengers into the markets in response to Northwest’s low fares. Removing
the high fares paid by these non-price-sensitive passengers left the fares paid by price-sensitive passengers.

Using this procedure, Spirit estimated the average fare that Northwest charged these non-price-sensitive passengers during the alleged predatory period in each market and adjusted them for commissions paid to travel agents as before. These values of Northwest's average net revenue per passenger are shown in Table 8-1. These revenues were $50.35 in the DTW-PHL market during July to September 1996 and $64.82 in the DTW-BOS market during April to September 1996. Comparing these revenues with Northwest's incremental costs shows that revenues fell short of incremental costs in both markets during the periods of alleged predatory pricing.

Spirit contended that the revenue-cost comparisons in Table 8-1 established that Northwest's pricing and capacity expansion measures were non-remunerative in the short run and created a presumption that Northwest's entry response in the two markets was predatory. Disputing the existence of a price-sensitive local passenger market, as noted earlier, Northwest maintained that Spirit's comparisons of revenues and costs for price-sensitive local passengers were contrived and inappropriate for detecting whether Northwest's response to Spirit's entry was predatory.

Northwest's Analysis of Revenues and Costs

The major difference between Spirit's and Northwest's analyses stemmed from their different perspectives on market definition. Northwest did not exclude connecting passengers from the relevant product markets, as Spirit did. Nor did Northwest exclude those local passengers whom Spirit designated as price-insensitive. Northwest's more expansive market definitions made comparing Northwest's revenues and costs significantly less complicated than did Spirit's analysis.

All-passenger market definitions prompted Northwest to include "behind and beyond contributions" of the connecting passengers on the revenue side of its analysis. These are the net revenues that Northwest attributes to a given route because they are earned on other legs of connecting passengers' flights. Since each Northwest flight in and out of DTW is part of a network, eliminating a connecting passenger on that flight would reduce the company's net revenues on other flights elsewhere in the network.

When connecting passengers were included in the product markets, together with their behind and beyond contribution to revenue, Northwest observed that it was no longer necessary to calculate and compare average variable costs to reach a conclusion. This is because Northwest's revenue exceeded even the carrier's fully allocated costs in the DTW-PHL and DTW-BOS markets in every month of the alleged predatory period. In fact, including connecting passengers in the product markets without including their behind and beyond contribution meant revenues exceeded fully allocated costs in every relevant month but May 1996 (in DTW-BOS). In addition,
setting aside behind and beyond revenues, Northwest's remaining all-passenger revenues in both markets always exceeded several incremental cost measures that Spirit relied on in its analysis.

For these reasons, Northwest maintained that the company’s entire operations on the DTW-PHL and DTW-BOS routes remained profitable on a fully allocated cost basis throughout the alleged predatory period. Northwest also maintained that, even ignoring behind and beyond revenue, its operations were profitable on any plausible incremental cost basis. Northwest concluded that these findings showed that the company’s response to Spirit’s entry on these routes was profitable in the short run and therefore nonpredatory.

Recoupment

Spirit’s analysis of Northwest’s fares and costs indicated that Northwest's fares were not remunerative in the short run and were unprofitable in the long run unless they served to drive Spirit from those markets, as Spirit contended that they did. Consequently, Spirit proceeded to the third and final step in the analysis: an assessment of the plausibility of Northwest's prospects for recouping the losses incurred in the DTW-PHL and DTW-BOS markets. Since Northwest’s analysis of revenues and costs indicated that the carrier had no loss to recoup, Northwest did not offer a detailed recoupment analysis. Northwest contended that this step was not necessary to discredit Spirit’s predatory claims because Northwest's revenue-cost analysis had already discredited them.

One can think of predatory pricing as being like an investment opportunity. This is because predatory pricing involves an incumbent firm charging prices that are below otherwise prevailing prices to impose a loss on an entrant. If the predator’s low prices are going to inflict a loss on an equally efficient entrant, the predator must incur a loss as well. The rational predator views this loss as an investment in a restored monopoly. The firm is willing to undertake this investment because it drives the entrant out of the market and subsequently pays an attractive return in the form of a monopoly profit during the hoped-for period of recoupment.

The insight that predatory pricing is an investment strategy is what lies behind the recoupment test. This test assesses the likelihood that the investment strategy implicit in the plaintiff’s predatory pricing complaint would be remunerative for the defendant, based on the postpredation market conditions that the defendant anticipates. To implement the recoupment test, one must calculate the loss that the alleged predator reasonably would expect from charging prices low enough and long enough to force the

42 This perspective is more fully explained and illustrated using data from several prominent antitrust cases in Elzinga and Mills (1989).
entrant out of the market. Then one must gauge whether the monopoly profit that the alleged predator reasonably would expect to earn after the entrant withdraws or is beaten back is large enough and lasts long enough for the putative predator to recoup this loss.

Recoupment should take into account all of the alleged predator's prospective benefits and costs from an episode of predatory pricing, regardless of whether those benefits and costs occur inside the relevant product market. Spirit's implementation of the recoupment test considered only the all-local-passenger product markets in DTW-PHL and DTW-BOS. It did not offer a separate recoupment analysis for the price-sensitive local passenger markets. In Spirit's all-local-passenger markets, the main benefits and costs of the alleged predation would occur within those markets. In the price-sensitive local-passenger markets, however, there may be benefits and costs that occur in adjacent markets. Spirit reasoned that the all-local-passenger markets analysis applied to the price-sensitive local-passenger markets as well because it would be improper to disregard those financial implications of Northwest's alleged predation that fall outside the price-sensitive markets.

To implement the recoupment test in the DTW-PHL and DTW-BOS markets, Spirit made some assumptions about what Northwest reasonably might have expected to happen when it launched its response to Spirit's entry. These assumptions imputed to Northwest the ability to foresee what its likely future fares and passenger loads would be, whether or not Northwest launched its allegedly predatory response. Specifically, Spirit assumed that Northwest could reasonably anticipate what its fares and the number of passengers carried would be during the alleged predatory period and beyond (i.e., once Spirit exited) and that those expectations were accurate in retrospect. Similarly, Spirit assumed that Northwest could reasonably anticipate what those values would have been had it not responded as alleged. In other words, Spirit assumed that Northwest, based on its costs and the fares and passenger loads under both scenarios, could compare its anticipated loss during predation and the gain it anticipated during subsequent recoupment. Spirit might have argued that Northwest expected that predatory pricing against Spirit in DTW-PHL and DTW-BOS would establish or strengthen Northwest's reputation for aggressive responses that would deter entrants (or capacity expansion by rivals) in other markets. But Spirit did not explicitly impute a reputation-building motive to Northwest or base its analysis of recoupment on such prospects.

Spirit used two approaches to estimate Northwest's otherwise prevailing fares. One (cost-based) approach estimated those fares using Northwest's fully allocated cost of service in the markets in question and

\[ \text{See the earlier discussion on reputation effects.} \]
imputing the same relationship between Northwest’s contemporaneous fares and costs that, on average, applied in all other DTW-based markets where conditions were similar and where there was no allegation of predation. To estimate the number of passengers that Northwest would have flown under the otherwise prevailing fares, Spirit extrapolated from actual passenger loads and fares using a range of plausible values for the price elasticity of demand for service. Northwest took issue with this approach because it relied on the relationship between Northwest’s fares and costs on all of its DTW-based routes rather than just those routes that are similar in length, number of competitors, and other factors. Northwest also contended that otherwise prevailing fares should not be projected from fully allocated costs in the first place.

The other approach that Spirit used to estimate Northwest’s otherwise prevailing fares exploited the fact that Northwest and Spirit also competed with each other on flights between DTW and New York’s LaGuardia Airport (LGA). This (comparable market) approach first projected Spirit’s fares in the DTW-PHL and DTW-BOS markets based on Spirit’s fares in DTW-PHL during the last three months of its single-flight operations (and before adding a second flight and precipitating Northwest’s alleged predatory response). Northwest’s otherwise prevailing fares in the markets in question were then estimated by imputing the same relationship between the two airline’s fares as the one observed during the same time period on the DTW-LGA route.7

Spirit used these estimates of Northwest’s (actual and hypothetical) fares and passenger loads to calculate the month-to-month loss and gain that Northwest might anticipate from predatory pricing. An annual discount rate of 15 percent was used to put these gains and losses on a comparable basis.8

On the assumption that the price elasticity of demand for service in these markets is -0.65, Table 8-2 indicates how many months of recoupment Northwest would require in order to break even on its investment in predatory episodes of two to four months in the DTW-PHL market and four to six months in the DTW-BOS market.9 These numbers of months of anticipated predatory pricing in the two markets bracket the number of months that actually passed between the onset of Northwest’s low prices

---

7 Corresponding estimates of the number of passengers that Northwest would have flown under these fares were constructed in the same way as before.

8 Ibbotson’s Cost of Capital Quarterly information for the months of March, June, September, and December 1996 for the airline industry (SIC Code 4512) assigns a range of 11.68 to 12.90 percent for the industry composite weighted average cost of capital. Based on this, using 15 percent kept the analysis on the conservative side.

9 Spirit used alternative estimates of the price elasticity of demand for service ranging from -0.3 to -1.0. These values were taken from the testimony of a Northwest executive. See Spirit Airlines Inc. v. Northwest Airlines Inc., 431 F.3d 917, 924 (2005). Table 8-2 assumes that the price elasticity of demand is the midpoint of this range.
and Spirit’s exit: three months in DTW-PHL and (almost) five months in DTW-BOS.

Based on these estimates, Spirit maintained that the entry (and reentry) barriers cited previously—especially the scarcity of gates at DTW and Northwest’s dominance of existing gates—meant that once Spirit withdrew from either market, Northwest would be unlikely to face another entrant in that market for much longer than the company needed to recoup its predatory sacrifice. As a result, it was reasonable for Northwest to expect that its alleged investment in predatory pricing would pay off. Spirit bolstered this conclusion by observing that, in retrospect, no significant entry took place in DTW-BOS for many years after Spirit withdrew from the market. The only significant entry event in DTW-PHL since Spirit withdrew from the market occurred in May 1998 when Pro Air introduced service between a secondary Detroit airport and PHL. This happened nineteen months after Spirit’s withdrawal, which left ample time for Northwest to recoup its investment in predatory pricing before Pro Air challenged Northwest’s dominance in the market.

Northwest disputed Spirit’s assessment of entry (and reentry) barriers that purportedly facilitated recoupment on the two routes. Citing the Pro Air experience and Spirit’s continuing operations to other cities at DTW, Northwest claimed that entry conditions in DTW-PHL and DTW-BOS would not provide the necessary shelter for Northwest to recoup alleged predatory losses. The carrier also contended that, barring evidence of collusion, competition from US Airways would prevent Northwest from recouping alleged predatory losses in the DTW-PHL market. Northwest cited two reasons why it might expect Spirit to remain in the DTW-PHL and DTW-BOS markets in spite of Northwest’s response to Spirit’s entry.

### TABLE 8-2
Northwest’s Months to Recoupment in DTW-BOS (months)

<table>
<thead>
<tr>
<th>Basis for Estimating Otherwise Prevailing Fares</th>
<th>DTW-BOS</th>
<th>DTW-PHL</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cost Based</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anticipated 2 Months Predation*</td>
<td>n.a.</td>
<td>1</td>
</tr>
<tr>
<td>Anticipated 4 Months Predation*</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Anticipated 6 Months Predation*</td>
<td>4</td>
<td>n.a.</td>
</tr>
<tr>
<td><strong>Comparable Market Based</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anticipated 2 Months Predation*</td>
<td>n.a.</td>
<td>3</td>
</tr>
<tr>
<td>Anticipated 4 Months Predation*</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Anticipated 6 Months Predation*</td>
<td>9</td>
<td>n.a.</td>
</tr>
</tbody>
</table>

Source: Northwest Airlines, Spirit Airlines, DB 1A data.

*These estimates assume that the price elasticity of demand for service in DTW-BOS and DTW-PHL is -0.65.

and Spirit’s exit: three months in DTW-PHL and (almost) five months in DTW-BOS.

Based on these estimates, Spirit maintained that the entry (and reentry) barriers cited previously—especially the scarcity of gates at DTW and Northwest’s dominance of existing gates—meant that once Spirit withdrew from either market, Northwest would be unlikely to face another entrant in that market for much longer than the company needed to recoup its predatory sacrifice. As a result, it was reasonable for Northwest to expect that its alleged investment in predatory pricing would pay off. Spirit bolstered this conclusion by observing that, in retrospect, no significant entry took place in DTW-BOS for many years after Spirit withdrew from the market. The only significant entry event in DTW-PHL since Spirit withdrew from the market occurred in May 1998 when Pro Air introduced service between a secondary Detroit airport and PHL. This happened nineteen months after Spirit’s withdrawal, which left ample time for Northwest to recoup its investment in predatory pricing before Pro Air challenged Northwest’s dominance in the market.

Northwest disputed Spirit’s assessment of entry (and reentry) barriers that purportedly facilitated recoupment on the two routes. Citing the Pro Air experience and Spirit’s continuing operations to other cities at DTW, Northwest claimed that entry conditions in DTW-PHL and DTW-BOS would not provide the necessary shelter for Northwest to recoup alleged predatory losses. The carrier also contended that, barring evidence of collusion, competition from US Airways would prevent Northwest from recouping alleged predatory losses in the DTW-PHL market. Northwest cited two reasons why it might expect Spirit to remain in the DTW-PHL and DTW-BOS markets in spite of Northwest’s response to Spirit’s entry.
First, Northwest reasonably might have assumed that, having entered the two markets with low fares, Spirit would remain in those markets charging those low fares. Second, if Spirit genuinely believed that Northwest’s fares were nonremunerative, the entrant would remain in the two markets until Northwest raised fares to remunerative levels.

CONCLUSION

So what happens next? Ever since the Supreme Court’s opinions in Matsushita (1986) and Brooke Group (1993), plaintiffs in predatory pricing cases have had a difficult time in court. The American Airlines case underscored that this difficulty extended to the airline industry. In Spirit v. Northwest, the district court’s initial opinion was in the Matsushita and Brooke Group grain: Northwest won on summary judgment.¹ However, on appeal, the Sixth Circuit Court of Appeals found that a reasonable trier of fact could find Spirit’s analysis of Northwest’s conduct credible and remanded the case to the district court for a full trial.²

Spirit v. Northwest affords a timely opportunity for courts to reconsider how the prudent Brooke Group approach to deciding predatory pricing cases should be applied in the airline industry. Unlike Matsushita and Brooke Group, in which alleged predatory pricing episodes did not cause the demise of efficient rivals followed by higher prices, the facts in Spirit v. Northwest feature the exit of a viable competitor and a subsequent increase in prices. Antitrust should not stand in the way of a new cohort of efficient producers displacing less efficient incumbents (Matsushita), or of incumbents introducing new products to compete with those pioneered by a rival (Brooke Group). Whether antitrust should stand in the way of an incumbent airline’s losing money by dropping its fares to match or undercut a low-fare entrant in a market and flooding the market with enough capacity to reduce the entrant’s load factors to unsustainable levels is a question that Spirit v. Northwest should decide.

Because airlines typically operate in many markets and because they can easily redeploy inputs from one market to another, the economic analysis required in a predatory pricing case in the airlines industry is complicated. Mastering the subtleties of airline market definition, price-cost analysis, and recoupment is a lot to ask of a trial court. In reference to Spirit v. Northwest, Kahn (2006, p. 174) recently wrote: “One can only hope that the jury to which the Circuit Court has consigned the case will be presented with the question in plain English of which side represented

---

preservation of the competitive process.” It is not clear that his hope for this case can be met.

Werden’s (2003) account of the American case suggests that this much-watched predatory pricing trial ended with a “whimper.”” Notwithstanding its potential importance for revising or endorsing the current ground rules that govern incumbent airlines’ response to entry, Spirit v. Northwest also may end with a whimper, but for a different reason. During the course of the litigation, Northwest, like several other major U.S. airlines, entered into bankruptcy. As of early 2007, this has stalled proceedings by preventing Spirit from returning to the courts for a final resolution.

REFERENCES


“See also Edlin and Farrell (2004).


Kahn, Alfred E. “Telecommunications: The Transition from Regulation to Antitrust.” Journal on Telecommunications and High Technology Law 5 (October 2006): 159-188.


United States v. AMR Corp. 335 F. 3d 1109 (10th Cir. 2003).


CASE 9


John P. Bigelow and Robert D. Willig*

INTRODUCTION

In 2001, the FTC brought a monopolization case against Schering-Plough, Upsher-Smith Laboratories, and ESI Lederle (a subsidiary of American Home Products), challenging agreements that settled the lawsuits that Schering had filed against Upsher-Smith and ESI for allegedly infringing Schering’s patent on its potassium chloride tablet, K-Dur 20. The FTC alleged that these agreements included so-called reverse payments from a patent holder to an accused patent-infringer as part of a deal between them to settle the patent litigation.

It was the view of the FTC that such reverse payments are a sure sign that the settlements were anticompetitive restraints of trade because, according to the FTC, they inherently buy protection for the patent holder against competition from the possible entry of the alleged infringer. The defendants argued that the only true reverse payments were directed by the judge in one of the patent cases, that reverse payments may be necessary to achieve socially beneficial settlements of patent litigation, and that there

*The authors served as consultants to Schering-Plough, and Willig gave expert testimony on behalf of Schering-Plough in the case described here.

‘FTC (2001a). Such payments are called “reverse payments” by those who are suspicious of them because they believe that the “natural” direction in which payments should flow in settling patent infringement litigation is from the alleged-infringer/defendant to the patentee/plaintiff, rather than the other way around.
was no showing by the FTC that competition was diminished by the patent settlements.

Over a period of more than five years the case wended its way before an FTC administrative law judge (ALJ), before the full Federal Trade Commission, and into the Eleventh Circuit Court of Appeals, prior finally to being considered for hearing by the Supreme Court. In 2005, when the Supreme Court declined to hear the case, the legal processes were exhausted, and Schering-Plough had prevailed over the FTC; but it was by no means clear that the substantive general issues raised by the case had been resolved.

BACKGROUND

The case arose in the context of competition between generic and brand name pharmaceuticals, a setting with specific institutional and regulatory features that are relevant to this case. When a pharmaceutical company has developed a new drug, it will generally patent the drug or its formulation or both. Before the drug can be marketed, the Food and Drug Administration (FDA) must be satisfied that it is safe and effective. To demonstrate safety and efficacy, the manufacturer submits a New Drug Application (NDA) to the FDA describing the results of the drug's clinical tests, its ingredients, the results of animal studies, the behavior of the drug in the body, and the drug's manufacturing and packaging. Upon approval of the NDA by the FDA, the manufacturer can begin selling the drug.

Generic drug manufacturers can, in effect, “piggy-back” on the NDA process. The manufacturer of a generic drug can rely on the safety and effectiveness conclusions of the NDA and need submit only an Abbreviated New Drug Application (ANDA), which provides the FDA with sufficient data to determine the generic’s bioequivalence (i.e., the ability to deliver the same amount of the drug into a patient’s bloodstream in the same amount of time) to an already approved drug. An ANDA generally does not need to include data from new animal and clinical studies.

The ANDA process was created by the Drug Price Competition and Patent Term Restoration Act of 1984, better known as the Hatch-Waxman Act, which both eased the process whereby generic drugs could win FDA approval and restored some of the patent life lost to manufacturers of new drugs during the drug approval process. This act established how the interaction should operate between manufacturers of brand name drugs, who generally hold patents, and the manufacturers of generic drugs. NDA filers must identify all the patents that they hold that pertain to their product. These are included in the drug’s listing in the FDA “Orange Book,” which is a list of all FDA-approved drugs.

For authoritative descriptions of the regulatory context, see FDA (2007a and 2007b).
When a generic drug manufacturer files an ANDA, it must certify "(I) that the required patent information has not been filed; (II) that the patent has expired; (III) that the patent has not expired, but will expire on a particular date and approval is sought after patent expiration; or (IV) that the patent is invalid or will not be infringed by the generic drug for which the ANDA applicant seeks approval." The last type of certification ("paragraph IV") gives rise to the kind of patent litigation involved in this case. The certification must include the ANDA filer's rationale for believing that the patent is invalid and is supplied not only to the FDA but also to the patent holder. The patent holder has forty-five days in which to file a patent infringement suit against the ANDA filer, and if the patent holder does so, then approval of the ANDA is stayed for thirty months or until either the patent expires or a final determination is made that the patent is invalid or not infringed. The first paragraph IV ANDA filer is allowed a period of 180 days of generic exclusivity, during which additional ANDAs will not be approved. The 180 days run from the first day that the generic is marketed, the patent expires, is found to be invalid, or is found not to be infringed by the ANDA.

The product at the center of the case was Schering-Plough's sustained-release potassium chloride tablet that is sold under the trade name K-Dur 20. It is taken to treat potassium deficiency, a common condition for the many patients taking medicine for high blood pressure. K-Dur 20 tablets were the most convenient form of the needed treatment and maximized patient compliance with doctors' orders. Annual sales of K-Dur 20 tablets were greater than $200 million at their peak. At the time that the case was brought, Schering held a patent (the "743 patent," as shorthand for Patent No. 4863743), expiring in September 2006, on the specific coating of K-Dur 20 that slowly releases the potassium chloride over time, making it a sustained-release product.

In 1995, Upsher-Smith Laboratories filed an ANDA for a generic version of K-Dur 20. As the first paragraph IV ANDA filer, Upsher-Smith would be entitled to the 180-day generic exclusivity period. Schering sued Upsher for infringing the 743 patent, and in June of 1997, just before trial was set to begin, the parties reached a settlement. Under the terms of the settlement, Upsher received a royalty-free license to market its generic product beginning in September 2001. Schering licensed five products, including Niacor SR, a sustained-release niacin product, from Upsher for which it paid royalties of $60 million.

The details of this certification process and a view of its impacts are well summarized in FTC (2002b), pp. 5-7.

There is substantial agreement among the parties concerning the basic facts in the context of the case that are summarized here. For the details, see FTC (2001b) and Schering-Plough Corporation (2001a and 2001b).
ESI-Lederle, a subsidiary of American Home Products, also filed a paragraph IV ANDA for a generic version of K-Dur 20 in 1995. Schering sued ESI for infringing the 743 patent, and the district court judge placed the parties under the supervision of a magistrate judge to work toward a settlement. Under the magistrate's encouragement a settlement was reached in January 1998 that allowed ESI to begin marketing its product in January 2004. ESI received $5 million from Schering initially and an additional $10 million that was contingent on ESFs receiving tentative approval of its product from the FDA by June 30, 1999. Additionally, Schering paid $15 million to license two unrelated products from ESI.

The FTC's Complaint Counsel (CC), the lawyers from the FTC who act as plaintiffs arguing the case against the respondents (defendants), formulated the complaint. They charged that both agreements unreasonably restrained trade, that Schering had attempted to monopolize the market for potassium chloride supplements, and that Schering had conspired with Upsher-Smith and ESI to monopolize the market. The payments from Schering-Plough to the other two firms occupied center stage in the complaint's allegations: "By making cash payments to Upsher-Smith and ESI, Schering induced them to agree to delay launching generic versions of K-Dur 20. Absent those payments, neither Upsher-Smith nor ESI would have agreed to delay its entry for so long."

As the case subsequently developed, three complex issues in law and economics were at its center:

1. Is a settlement of a patent infringement case that embodies a reverse payment intrinsically anticompetitive? What might be pro-competitive reasons for reverse payments? Should reverse payments be subject to a per se rule or to rule-of-reason treatment?

2. Is it impossible or prohibitively expensive for an antitrust inquiry to evaluate the likelihood of the parties' prevailing in the underlying patent litigation?

3. What is the appropriate standard to use in evaluating whether or not the outcome of a settlement of a patent infringement case is or is not anticompetitive?

The three questions are related, especially from the perspective of a policymaker. If the right way to determine if a settlement is anticompetitive is to compare the outcome of the settlement with the expected outcome of the underlying litigation, then one would need to assess the likelihood of

*Tentative approval is given to ANDA filers when the requirements for approval are met, but final approval is stayed because of a patent.

Federal Trade Commission (2001a), para. 64.
alternative outcomes of that litigation to reach a conclusion about the settlement. If that is impossible or impractical, a short-cut rule such as a blanket condemnation of settlements with reverse payments may seem attractive. But such a per se rule might be a significant impediment to settlements of patent litigation, and there is a general view that litigation settlements are socially efficient in that they conserve scarce judicial resources, save on substantial litigation costs, and eliminate otherwise unnecessary risk-bearing by the litigants.

Further, it may be the case that reverse payments can enable pro-competitive settlements of patent litigation with sufficient frequency that per se condemnation of them would cost society more than it would gain by cutting off those settlements with reverse payments that are truly anticompetitive. Thus, it may be the case that a rule-of-reason approach would be best, wherein a weighing of factors, perhaps including the size of the reverse payment, would be permitted to decide the matter on a case-by-case basis.

In what follows, we trace the history of this case, the arguments made, and the conclusions drawn with respect to these fundamental issues.¹

THE TRIAL

The case was first tried before an administrative law judge (ALJ) within the FTC. That trial commenced on January 23, 2002, and ended on March 28, 2002. It filled 8629 pages of transcript, with forty-one witnesses testifying and thousands of exhibits admitted into evidence. Closing arguments were heard on May 1, 2002.²

Reverse Payments

Entry, Litigation, and Expected Entry

While the economists who testified disagreed in their conclusions about the competitiveness of reverse payments, there were aspects of their analyses upon which they agreed. The economists for both sides viewed the patent litigation as an event with a random outcome. With some probability the incumbent/patentee would win, and entry would not take place until the end of the life of the patent; and with some probability the entrant/generic manufacturer would win, and entry would take place sooner.

¹These issues have engendered a significant outpouring of professional literature. For a more complete and technical articulation of our views, see Willig and Bigelow (2004). For a focus on the social benefits of settlements, see Langenfeld and Li (2003). For a more comprehensive bibliography on the instant issues, see Cotter (2004). For application of some rigorous economic tools to these issues in a delimited setting, see Shapiro (2003).
The economists for both sides compared the date upon which the settlement agreement called for entry with the expected date of entry pursuant to the litigation. The latter is an entry date that would be realized if the life of the patent after the litigation was concluded were divided proportionally to the probabilities of the parties’ prevailing in the litigation. For example, suppose that the remaining life of the patent is twelve years, the course of the litigation will consume two years, and that the probability is 60 percent that the incumbent would win the patent litigation. Eight years from now is the time of entry that would be consistent with two years of exclusivity during litigation and a 60 percent chance of continuing exclusivity for the remaining ten years of the patent’s life. The four years of nonexclusivity under the settlement are consistent with the entrant’s 40 percent chance of being permitted to compete for ten years following the two years of litigation.

If the demand for the product and the cost of producing it do not change throughout the life of the patent, if the parties were risk neutral, if the parties did not discount the future, and if the litigation were costless, then each of the parties would be indifferent between the litigation and a settlement that provided for entry on the expected date of entry under litigation. To the extent that the parties are risk averse or the litigation is costly, then each of the parties would prefer a settlement with an entry date somewhat less favorable than the expected date of entry under litigation; that is, the incumbent would be willing to accept an earlier date, and the entrant would be willing to accept a later date, to the litigation itself. Likewise, if consumers are risk neutral, any settlement with an entry date before the expected date of entry under litigation is better for consumers than the litigation. If consumers are risk averse, then even some settlement dates after the expected date of entry under litigation are preferred.

Complaint Counsel’s Case against Reverse Payments

In order to make its case that Schering had made reverse payments to the two generic manufacturers in exchange for entry delay, CC first had to identify the reverse payments. In the ESI settlement, CC identified the $5 million initial payment and the $10 million payment that was contingent on ESI’s winning approval from the FDA, drawing special attention to the fact that the later that ESI (AHP) had an approvable product, the less money it would receive. In the Upsher-Smith settlement, CC argued that the $60 million payment was not a bona fide license fee and was, instead, a payment for delay. In support of this contention, CC argued that the $60 million payment to Upsher-Smith was “anomalous” in that it was noncontingent (i.e., not tied to performance on Upsher-Smith’s part), excessive relative to

More formally, the expected entry date is, in the language of probability, an expected value. Under litigation, the date of generic entry is a random variable. The expected date of entry is the sum of the products of the possible entry dates and their respective probabilities.
payments that Schering made for other licenses, and unsupported by Schering’s customary level of due diligence. CC also observed that Schering did not follow up by attempting to obtain regulatory approval for Niacor SR or begin work on manufacturing and marketing the product.

Turning to the economics of reverse payments, CC introduced an economist expert witness who argued as a matter of economic theory that a settlement with a reverse payment must be anticompetitive. The argument rested on two legs. First, he argued (in his expert report) that when confronted with the opportunity to strike an anticompetitive agreement, the firms would necessarily do so. He argued that if it were possible for the incumbent firm to make a payment to the entrant, the firms would never agree to a settlement that provided consumers with the same expected competition as litigation. To do so, he argued would, in effect, be to “leave money on the table.”

Second, he analyzed the role of reverse payments in the context of a simple model of the bargain between the incumbent/patentee and the entrant/generic. The assumed structure of that model included only the potential gains to the two firms from the extension of monopoly power and the redistributive (as between the two firms) capacity of payments from firm to firm. In the context of that model he concluded in his report that if the payment received by the entrant is essential to the entrant’s being willing to agree to a settlement, then the settlement must reduce competition, and the payment must be compensation for the profits that the entrant foregoes by agreeing to a shorter period of competition. In effect, he argued that the only reason that the incumbent would be willing to pay the entrant would be in exchange for delayed entry, and the only reason that the entrant would be willing to delay entry would be in exchange for a payment.

These summarized arguments can be made clearer in the context of a basic economic model of the parties’ incentives to settle a patent case. The two firms are the incumbent (firm I) and the would-be entrant (firm E). The potential economic life of the patent runs from the current time, \( t = 0 \) through \( t = 6 > 0 \). The rate of profit of the incumbent as long as it is a monopoly is \( a \) per unit of time. The rate of profit of the incumbent as a duopolist after the would-be entrant actually does enter is given by \( B \). The rate of profit of the entrant is 0 before it enters and is \( B \) per unit of time after it enters. Let \( p \), where \( 0 < p < 1 \), be the probability that the incumbent wins the litigation, and let \( C_I > 0 \) and \( C_E > 0 \) be the costs of litigation to the incumbent and the entrant, respectively. These represent the expected cash costs of the litigation plus the impacts of the risks of the litigation due to any risk aversion on the part of the litigants, expressed in money terms as risk premiums.

“For a classic reference on the incentives for litigation settlement, see Priest and Klein (1984). The development in the text is similar to that employed by both parties’ economic experts.”
Thus, the expected payoffs (profits) to the incumbent and the entrant from litigation are, respectively,

$$\bar{Y}^I(0) = p\alpha \theta + (1 - p)\beta \theta - C_i \quad \text{and} \quad \bar{Y}^E(0) = (1 - p)\beta \theta - C_E.$$ 

Let $t$ denote the agreed-upon date of entry, at some point between the current time (0) and the end of the potential economic life of the patent (0), pursuant to a settlement, and suppose that the incumbent firm makes a net transfer ("reverse") payment of $B$ to the entrant. The expected payoffs (profits) to the two firms under such an agreement are

$$\bar{Y}^I(t; B|0) = \alpha t + (\theta - \beta)B_2 - B = \theta B_2 + (\alpha - \beta) t - B \quad \text{and} \quad \bar{Y}^E(t; B|0) = (\theta - \beta)B_2 + B.$$ 

A settlement involving the agreed-upon date of entry $t$, together with the reverse payment $B$, will be acceptable to both parties if, for each of them, the expected payoff under the settlement is greater or equal to the payoff from litigation. Thus, with a rearrangement of the resulting inequalities to focus on $t$, a settlement is acceptable to both parties if the following holds:

$$p\theta + \frac{B - C_i}{\alpha - \beta} \leq t \leq p\theta + \frac{B + C_E}{\beta \theta}.$$

It is key to recognize that $a > \beta \theta + /3$. This inequality expresses the likely fact that the flow of monopoly profits to the incumbent when it is the only seller of the product at issue is greater than the sum of the flows of the profits to both the incumbent and the entrant when they are competing as duopolists selling the product. This is generally true inasmuch as the monopolist chooses its output level and price to maximize profit, while the duopolists are in part driven in their decisions by their competition against each other, which tends to lower profits even as it likely benefits consumers of the product.

In this typical circumstance, there is a range of values of $t$ that enables settlement on mutually beneficial terms, since with $a = \beta \theta + /3$, the left-hand side of the inequality must be less than the right-hand side. Without any reverse payments, $B = 0$, and there will be mutually agreeable entry dates $t$ so long as there are costs of litigation and risk aversion. In fact, with positive litigation costs there will be mutually acceptable settlements that entail entry at an earlier date than $pd$, which is the expected date of entry under litigation. There will also be later entry dates that are mutually agreeable.

However, with reverse payments, as $B$ is positive and continues to grow bigger, the dates $t$ that are mutually acceptable grow larger also, signifying
increasingly later entry and more time of monopoly over the product. The intuition is that a later entry date \( t \) substitutes more monopoly time for more duopoly time and thereby creates a larger total flow of profits. This result is again due to the key relationship that \( a > B + B \). The incumbent can share the greater flow of profit with the entrant by means of the reverse payment; and, while the parties are better off, consumers are worse off due to the suppression of competition. In the limit, the entry date can be pushed back to the end of the patent’s useful life, \( 0 \), and an appropriate reverse payment can make both incumbent and entrant better off than they would have been with litigation. This profit-maximizing arrangement eliminates any possibility of competition and so is both best for the parties and worst for consumers.

In essence, it is this set of analytic arguments that was articulated in the economic testimony of the expert engaged by CC, and it was the logical underpinning of the case brought by CC against Schering-Plough. Moreover, these arguments underlie the FTC’s decision to attack reverse payments on a per se or presumptive basis. In the basic model just presented, no good can come from reverse payments because they are not necessary to encourage settlements, they move settlements only in the one direction that suppresses competition; and it is most profitable for the patent litigants to employ reverse payments to eliminate as much competition as they can.

Respondents’ Case in Favor of Reverse Payments

Schering’s expert witness disagreed with these conclusions as a matter of economic theory. Schering’s expert argued that CC’s expert had gotten the economics of reverse payments wrong by employing an economic model that was too simple—in effect assuming away the features that give rise to the pro-competitive features of reverse payments. The gist of the critique was that by using a model in which entry delay and monetary payments were the only two aspects of the settlement available to the parties to bar gain over, one inevitably emerged as a quid pro quo for the other. By contrast, in a richer and more realistic model, payments from the incumbent to the entrant can be consistent with pro-competitive settlements. For example, litigation is costly, and a settlement (if it doesn’t lead to antitrust litigation) is not. An incumbent who is risk averse would be willing to pay to avoid risk. So, a risk-averse incumbent would gain by paying an entrant to agree to a settlement that called for entry on the expected date of entry under litigation.

More importantly, Schering’s expert identified circumstances under which reverse payments are essential to achieving pro-competitive settlements. Under such circumstances, by using reverse payments the parties would be able to achieve mutually beneficial settlements that were better for consumers than litigation would have been. If the parties were forbidden to use reverse payments, then no settlement would be possible. These
situations were all excluded by the excessively simple model articulated above but are all quite intuitive and plausible. They are explained in the subsections that follow.

**Differences in Discount Rates: Cash-Strapped Entrant**

If the entrant firm has a general need for early cash but is unable to secure enough new credit from financial markets, it will discount the future very substantially relative to the present—probably much more so than does the incumbent firm. Litigation yields some positive probability of early cash flow resulting from a win in court, and such a firm would agree only to a settlement with entry at a very early date. However, such an early date would be unacceptable to an incumbent that had a lower discount rate and that believed it had a substantial probability of prevailing in the litigation. Absent the ability to use reverse payments, these two firms will be unable to agree on a settlement, and the (costly) litigation will inescapably proceed. However, the parties could agree on a mutually beneficial settlement with entry before the expected date of entry under litigation if the incumbent were risk averse and was able to make a reverse payment to the entrant to satisfy the latter’s need for immediate cash. Both firms and consumers gain from this settlement.

**Information Asymmetry**

In many cases the incumbent will know more about the value of the market than does the entrant. For example, the incumbent, because it is involved in research and development, may know more about which drugs that could potentially displace the current product are in the pipeline. The rational entrant, knowing that the incumbent has this information that the entrant does not have, must interpret the incumbent’s actions with a view toward the inferences that may be drawn from them about the incumbent’s information.

Such a process of inference can cause settlement negotiations to break down. The phenomenon at work here is a familiar one. If I offer to sell a late-model used car for an exceptionally low price, the wary buyer’s first reaction is likely to be, “What’s wrong with it?” The prospective buyer, knowing that I have information about the condition of the car, reasons that I ought not to be willing to sell a high-value car for a very low price, interprets my willingness to sell the car as an indication that it is of low value, and may decline to buy the car.\(^1\)

To see the consequences of this process in the context of patent litigation settlement, suppose that the value of the patent is either high or low and

\(^1\)This is the familiar “lemons” or adverse selection problem; see, e.g., Akerlof (1970).
that the incumbent knows which it is, but the entrant does not. Also suppose that when the value of the patent is low, the cost of litigation is too high relative to the value of the patent for litigation to be worthwhile to the incumbent, so if a settlement is not reached the incumbent will concede the result of the lawsuit to the entrant.

Suppose initially that the parties are not permitted to employ a reverse payment in a settlement. A settlement may be impossible. In this setting the incumbent who knows that the value of the patent is low will agree to any settlement that would be acceptable when the value of the patent is high. In other words, there are only two possible types of settlements: those that both types of incumbents will participate in and those that only the low type of incumbent will participate in.

If the entrant expects the former, then it will evaluate possible entry dates on the basis of the probabilities it attaches to confronting a high or low type incumbent. In other words, the entrant must hedge the entry date to which it will agree to take into account the possibility that the value of the patent is low. That assessment may leave the entrant unwilling to agree to an entry date late enough to be acceptable to an incumbent that knows that the value of the patent is high. In that case, there can be no settlement involving both high and low type incumbents. On the other hand, if the entrant expects that the only incumbent that will agree to a settlement is an incumbent that knows that the value of the patent is low, then the entrant has no incentive to agree to a settlement. An incumbent that knows that the value of the patent is low will concede the litigation, allowing the entrant to enter right away. In these circumstances, therefore, no settlement without a reverse payment is possible.

On the other hand, if reverse payments are possible, then an incumbent that knows that the value of the patent is high will be willing to make a cash payment to an entrant in connection with a settlement that saves the incumbent its costs of litigation. The willingness of the incumbent to make a payment to the entrant signals credibly to the entrant that the incumbent knows that the value of the patent is high because an incumbent that knew that the value was low would be unwilling to make such a payment. With this information in hand, the parties can reach an agreement that is mutually beneficial and calls for entry before the expected entry date under litigation, thereby making it beneficial to consumers as well.

A numerical example may help to shed light on this circumstance: The incumbent has private knowledge of the useful life of the patent, and the entrant is aware of that and believes that the life is either six years with probability 0.33, or one year with probability 0.67 (after which entry by producers of similar drugs renders the patent valueless). The flow of monopoly profit is $3 per year, and the flow of profits to either duopolist is just $1 (and it is $0 if the patent is valueless). The patent litigation would go to the incumbent or the entrant with probability 0.5, but the incumbent faces litigation costs of $3, while the entrant faces no litigation costs at all.
Under these assumptions, if the incumbent knew that the patent were weak with just a one-year life, it would not be willing to litigate in view of its litigation costs, since its return from one year as a duopolist would be \(0.5(\$3) + 0.5(\$1) - \$3 < 0\). On the other hand, if the incumbent knew that the patent were strong with a six-year life, it would be willing to settle without any reverse payments only if the agreed entry date were at or later than 1.5 years (since profit from litigation = \((0.5)(\$3)(6) + (0.5)(\$1)(6) = \$3 = \text{profit from settlement} = (1.5)(\$3) + (4.5)(\$1)).

However, the entrant will not accept a settlement offer of entry rights at 1.5 years or later. Recall that the entrant thinks that with probability 0.67 the patent is weak and that with probability 0.33 the patent is strong. Consequently, agreeing to a settlement with entry rights at 1.5 years out would yield the entrant an expected profit of \((0.33)(6-1.5)(\$1) + (0.67)(0) = \$1.5\). Here, litigation would yield the entrant \((0.33)(6)(\$1) + 0.67(0) + (0.33)(0) = \$1.67\), since the incumbent will not contest the litigation when the patent is weak (and the entrant knows this). Thus, there is no settlement without reverse payments in the situation where the incumbent has private information that the patent is strong because the entrant is worried that the patent is really weak and the settlement would be worthless in that event.

Here a reverse payment can serve as a credible signal to the entrant that the patent is really strong and thus break the impasse that would otherwise make a settlement impossible. Consider the settlement that the incumbent would offer if the patent were strong—with entry allowed after 2.75 years and a reverse payment of \$2.25. On seeing this offer, the entrant would infer that the patent must be strong since if the patent were weak, the incumbent’s return from the offer would be one year of monopoly profit of \$3, less the reverse payment of \$2.25, which is less than what the incumbent would make by ceding the patent litigation and earning one year of duopoly profit of \$1.

With the inferred knowledge that the patent is strong, the entrant would accept the offer. This follows because the settlement would earn the entrant \((6 - 2.75)(\$1) + \$2.25 = \$5.5\), while litigating would earn \((0.5)(6)(\$1) + (0.5)(0) = \$3\). And the incumbent, knowing that the patent is strong, would prefer the settlement to litigation, since the settlement would earn it \((2.75)(\$3) + (6 - 2.75)(\$1) - \$2.25 = \$9.25\), while litigation would return \((0.5)(6)(\$3) + (0.5)(0)(\$1) - \$3 = \$9\).

In this example of a simple instance of asymmetric information, the employment of reverse payments makes a settlement possible where it would otherwise be impossible. Here, as in a rich set of other such examples, among the settlements made possible are Pareto improvements over anything that can be attained without reverse payments since the agreed entry date is earlier than the expected date under litigation. (In this example the expected date of entry under litigation is three years, while the date
of entry under the settlement is 2.75 years, all given that the patent is strong. If the patent is weak, then there will be no settlement and a duopoly for one year.)

**Varied Assessments of Success: Overoptimism**

A third set of circumstances under which reverse payments may be essential to pro-competitive settlements arises when the parties to the litigation hold different assessments of their chances of success. When the probability that the incumbent believes it will prevail and the probability that the entrant believes it will prevail sum to more than 100 percent, the incumbent's estimate of when the expected date of entry under litigation will occur will be later than the entrant's estimate. Under these circumstances it may be impossible for the parties to agree on a settlement entry date that is compatible with their respective expectations.

However, if reverse payments are permitted, negotiations become easier. This conclusion may be seen in a simple diagram. Consider Figure 9-1, in which the entry date in the settlement is measured on the horizontal axis, and the size of the reverse payment is measured along the vertical axis. The points on the diagram represent different possible settlements. If no reverse payments are allowed, then the only possible settlements are on the horizontal axis.

On the assumption that the firms are risk neutral and there are no litigation costs, continuing the litigation is equivalent to a settlement with no reverse payment and entry at the expected date of entry under litigation. If the parties vary in their assessment of success in the manner described above, then the incumbent's expected litigation entry date is later than the entrant's. These points are plotted on the horizontal axis.

**FIGURE 9-1 Settlement with Varied Assessment of Success**
We draw an indifference curve through the expected entry litigation entry date showing all the combinations of entry dates and reverse payments that are indifferent to the litigation for each firm. Both lines slope upward. The line for the incumbent slopes upward because a later entry date is needed to compensate the incumbent for making a larger reverse payment, and the line for the entrant slopes upward because the entrant is willing to accept a later entry date if it receives a larger reverse payment. The incumbent's line slopes up faster because each month of delay is worth more to the incumbent than it is to the entrant. Each additional month's delay brings the incumbent an increase in profit equal to the difference between monopoly profit and one firm's duopoly profit. Each additional month's delay costs the entrant one firm's duopoly profit. Since monopoly profit exceeds the sum of both firm's duopoly profits, the former is larger than the latter.

The incumbent would be willing to agree to any settlement on or below the incumbent's indifference curve (labeled "Incumbent's Litigation Line" in the diagram). The entrant would be willing to agree to any settlement on or above the entrant's indifference curve (labeled "Entrant's Litigation Line" in the diagram). Any settlement below the incumbent's line and above the entrant's line is agreeable to both firms. The fact that there are no mutually agreeable settlements on the horizontal axis reflects the impossibility of a settlement without a reverse payment.

The diagram also shows the actual expected litigation entry date. That expected entry date is based on the actual probabilities of the firms' prevailing, which may be different from the perceptions of either firm. On the assumption that consumers are also risk neutral, their indifference curve is a vertical line. The vertical line starting at the actual expected litigation entry date shows all the settlements that are indifferent to litigation for consumers. Consumers prefer settlements to the left of that line. When, as illustrated in Figure 9-1, the actual expected entry date occurs after the expected litigation dates of the two firms, there can be settlements, like those shown in the shaded region of the diagram, that are mutually agreeable to the two firms and are preferred to litigation by consumers. None of these settlements is possible without reverse payments.

Other Circumstances

The foregoing are three examples of circumstances under which reverse payments can enable the parties to achieve mutually agreeable settlements that are beneficial to consumers and are not possible without reverse payments. In addition to these, Schering's economist also described additional examples in circumstances of follow-on entry by additional generic manufacturers and the generic exclusivity provisions of the Hatch-Waxman Act.¹

¹These and similar circumstances are also described in Willig and Bigelow (2004).
In summing up this section, it should be recognized that the circumstances described are all examples where reverse payments may be necessary to effectuate socially desirable settlements. At the same time, in all of these circumstances, reverse payments could also be used as an element of settlements that support the extension of monopoly beyond what the patent litigation would otherwise yield in expected value. Thus, reverse payments are powerful tools that can be used to support anticompetitive as well as pro-competitive agreements in the sets of circumstances described here.

**Per se, the Rule of Reason, and Shifting the Burden of Proof**

In its trial brief, CC urged the ALJ to treat the reverse payments as per se unreasonable restraints of trade. “Because the nature of the agreements gives rise to an obvious inference of anticompetitive effects, they are unlawful per se unless they have a plausible justification. Respondents offer two purported justifications, but neither is plausible.” The basis for the inference referred to here is the presence of the reverse payment. CC leaves the door open in principle for the settlements to be defended upon the showing of “plausible justification” but in so doing would shift the burden of proof to the companies.

If the ALJ would not treat the settlements under the per se rule, CC was prepared to argue that they were summarily illegal under the rule of reason: “In the absence of plausible justifications, the agreements can be condemned as per se illegal. And even if per se treatment is not indicated, the Supreme Court made clear in NCAA that inherently anticompetitive restraints can be condemned under the rule of reason, if upon examination, the offered justifications are found to be inadequate. Under these standards the agreements can be summarily condemned.” Once again the CC’s approach would effectively shift the burden of proof to the firms to demonstrate that the agreements could be “justified.”

CC defended this shift by arguing that it had made a prima facie case against the settlements. “Once plaintiffs make a prima facie showing that an agreement is anticompetitive, it is defendants’ burden to come forward with a plausible procompetitive justification.” What is the basis of this showing?

CC conceded that it was relying on an inference that the settlements were anticompetitive. CC’s expert witness testified that as a matter of economic theory, settlements with reverse payments were necessarily anticompetitive. That would support the inference, but Schering’s expert had testified that CC’s expert had the theoretical conclusion wrong—that it was not true as a

---

14 FTC (2002a), p. 44.
matter of economic theory that a settlement with a reverse payment was necessarily anticompetitive. CC preferred to relegate this evidence to a second stage of the analysis, that of attempting to justify an already presumptively anticompetitive agreement. At that stage CC dismissed the arguments on the grounds that Schering's witness did not offer evidence that the circumstances he described pertained to the specific agreements in question. However, if CC’s prima facie case rested on the proposition that settlements with reverse payments are necessarily anticompetitive, and if that conclusion is wrong, then the second stage of the analysis should never have been reached because the argument failed at the first stage.

Evaluating the Underlying Patent Litigation

Schering's expert witness concluded from his analysis of the economic theory of reverse payments that evaluating the underlying patent so as to compare the settlement entry date with the expected date of entry under litigation was necessary to determine if a settlement was anticompetitive. Since settlements with reverse payments can be either pro-competitive or anticompetitive, examining settlements for the presence of a reverse payment will not distinguish pro-competitive from anticompetitive agreements.

CC disagreed, arguing that determining the probabilities of the outcomes of the underlying litigation is not only unnecessary but also impossible: "We will never know who would have won the patent cases. Still less can we quantify each party’s chance of success. There simply is no known methodology for handicapping trials or for testing the reliability of predictions." Additionally, CC expressed concern that, "by advocating tests that are impossible to apply they (Respondents) adopt a rule that would effectively immunize such settlements. Given the undeniable incentives for branded drug manufacturers and potential generic entrants to reach patent settlements that involve payments for delay, consumers are far better served by an analysis that places the burden on defendants to show that a reverse payment was not for delay, than one that would permit reverse payments for such settlements as a matter of course."

This argument also serves to explain why CC sought to shift the burden of proof to the respondents.

THE INITIAL DECISION

Five months after the trial began, on June 27, 2002, the ALJ handed down his decision ("The Initial Decision"). Except for some procedural questions

---

(such as finding that the case lay within the jurisdiction of the FTC, that the challenged practices impinged on interstate commerce, and that Schering-Plough and Upsher-Smith are corporations within the meaning of the FTC Act), the ALJ ruled against CC in virtually every substantive respect. He found that CC had failed to prove or define a relevant product market; that the agreements between Schering and Upsher-Smith and between Schering and ESI did not unreasonably restrain competition and were not unfair methods of trade; that Schering did not have monopoly power in the relevant product market, which the ALJ found to be “all oral potassium supplements that can be prescribed by a physician for a patient in need of a potassium supplement”; that Schering did not unlawfully engage in conduct to preserve monopoly power in the relevant product market; that Schering did not conspire unlawfully with Upsher-Smith or ESI to preserve monopoly power in the relevant product market; and that CC failed to meet its burden of proof.”

The Economics of Reverse Payments

The ALJ did not find it necessary to decide whether reverse payments are or are not intrinsically anticompetitive. First, he rejected the application of a per se rule on the grounds that patent splitting and reverse payments were issues that were too novel for application of the per se standard.” In rejecting the per se rule, he also touched on the question of the appropriate standard of comparison for anticompetitiveness. He concluded that because it is not clear that in the absence of the agreements the generics would have entered earlier than the entry dates in the settlements, the agreements cannot be condemned per se.” He also found that Schering’s conduct did not extend beyond the rights conferred by the patent. Similarly, the ALJ rejected a so-called quick-look analysis because he found pro-competitive effects sufficiently “plausible” that the quick-look approach was not appropriate.

Turning to a full-fledged rule-of-reason analysis, the ALJ concluded that the CC had not proved anticompetitive effects because it had not proved that a better settlement or litigation would have resulted in entry before the agreed-upon dates. “This does not necessarily mean that the ALJ rejected the logic of CC’s (and its witness’) argument that reverse payments are anticompetitive. Instead he rejected, in effect, the contention that CC had proved that a reverse payment took place. The ALJ concluded that CC had failed to demonstrate that the value of the Niacor license that Schering obtained from Upsher-Smith was not commensurate with the $60 million

"FTC (2002c), pp. 299a-301a.
"FTC (2002c), pp. 303a-308a.
that Schering paid for it.” The ALJ cited CC’s own expert witness for the proposition that the payment to Upsher-Smith was not for delay if it represented fair value for the license. “With respect to the payment to ESI, the ALJ criticized CC’s expert witness for relying on “what he [the witness] characterized as an ‘assumption’ that if ESI had won its patent suit, it might have been able to enter before March 2002.” Based on these findings, the ALJ concluded that the challenged settlements were of a kind that both CC and its witness had conceded were permissible.”

Evaluating the Patent

In contrast to his views on reverse payments, on the subject of evaluating the patent the ALJ agreed with CC. He concluded that, “There is no way to determine the date or the outcome of the judicial determination of the patent litigation.” However, notwithstanding this finding, the ALJ was not willing to shift the burden of proof to the respondents. Responding to the CC’s argument that the respondents’ proposed rules would “immunize” anticompetitive settlements, the ALJ decided, “simply because, based upon the theories it advanced in this case, Complaint Counsel cannot prove whether Upsher-Smith and ESI would have come on the market earlier than September 2001 and January 2004, but for the $50 million and $15 million payments, does not relieve Complaint Counsel of its burden of proof.”

The Standard for Anticompetitiveness

The ALJ held that in order to prove that the settlements were anticompetitive, CC had to define a relevant market and show that Schering had market power in that market. The ALJ rejected CC’s arguments that both steps were unnecessary and concluded that without this foundation, CC could not prove an anticompetitive effect.” The ALJ rejected as “circular” the argument that Schering must have had market power or it would not have paid Upsher-Smith and ESI not to compete. “The ALJ concluded that CC had not proved that Upsher-Smith and ESI had been paid to stay off the market because, as CC acknowledged, CC could not even show that the two generics could have been on the market before the expiration of the patent.”

THE COMMISSION’S DECISION

CC appealed the ALJ’s decision to the full Commission, which, unlike appeals courts accepting cases from lower courts, hears appeals from decisions of ALJ’s “de novo,” which means that the Commission can—and in this case did—review the record and make new findings of fact to replace those of the ALJ. (By contrast, generally, an appeals court is limited to reviewing the lower courts’ interpretation and application of the law to the facts that the lower court has found.) The Commission noted that while it had pursued a number of cases involving generics, this was the first one to come before the full Commission after a full trial and with a complete record.

In December of 2003 the Commission made public its opinion, which reversed the ALJ, finding that the agreements were unreasonable restraints of trade and imposing an order prohibiting the parties from being party to an agreement to settle a patent infringement claim in which an ANDA filer both receives something of value (beyond up to $2 million in legal fees) and agrees not to market or develop the ANDA product for some period of time. The order also forbade the parties from being party to an agreement in which an ANDA filer agrees to forgo selling or developing a product that is not subject to a patent claim.

Reverse Payments

The Commission disagreed with the ALJ that the licenses that Schering received were adequate consideration for the payments it made to the two generics. Consequently, the Commission rejected the ALJ’s conclusion that the adequacy of compensation meant that the payments were not for delay. In fact, the Commission concluded that the payments were anticompetitive. The Commission denied that it was shifting the burden of proving that the settlements were not anticompetitive to the respondents and attributed the burden of proving that they are anticompetitive to the prosecutor. However, the Commission apparently found that the following logic satisfied that burden: The payments must have had some offsetting consideration. In the absence of proof that the offsetting consideration was something else, it is "logical to conclude that the quid pro quo was entry delay." Employing this logic, the Commission was willing to conclude that the settlements had an anticompetitive effect, so it was not surprising that it also ruled that CC need not have defined a relevant market.

"FTC (2003a and 2003b)."
"FTC (2003b), pp. 76a-77a."
Applying the Rule of Reason and Fixing the Burden of Proof

Notwithstanding its description of the burden of proof, the Commission followed the line of CC and interpreted Schering’s expert’s critique of the argument that settlements with reverse payments are necessarily anticompetitive as an “affirmative justification” for the settlements rather than an attack on the CC’s prima facie case. The Commission evaluated the testimony of Schering’s expert as an “ancillarity claim”—in other words, as a claim that the agreement to delay entry (the ancillary agreement) was necessary to achieve the pro-competitive benefits of the agreement to settle the patent litigation (the main agreement).

Within this framework, the Commission could and did accept the testimony of Schering’s witness as correct without concluding that it undermined CC’s case against the settlements:

We also recognize, as he (Schering’s expert) testified, that there may be hypothetical situations where a procompetitive settlement could require payment of some money to the generic challenger. This means that we are unwilling to say reverse payments included in a settlement agreement are always illegal. On the other hand, the mere articulation of hypothetical circumstances where reverse payments could ultimately facilitate an efficiency-enhancing settlement does not mean that a particular settlement is legal. If Complaint Counsel have made out a prima facie case that the agreement was anticompetitive, the burden is on these Respondents to demonstrate that these hypothetical circumstances describe the realities of the present case. They have not done so.”

This line of reasoning, however, involved a paradoxical interpretation of the testimony of Schering’s expert, since it assumed that CC had made out its prima facie case, whereas that case relied fundamentally on the very proposition that Schering’s expert challenged.

Evaluating the Patent Litigation

The Commission concluded that it was possible to determine whether the settlements were likely to have had anticompetitive effects without assessing the merits of the patent litigation. In reaching this conclusion, the Commission rejected the relevance of the presumption that Schering’s patent was valid until proved invalid, pointing out that the issue in the patent litigation with Upsher-Smith was whether or not Upsher’s product infringed Schering’s patent—a question upon which Schering bore the burden of proof in the patent litigation. The Commission also drew a

"FTC (2003b), p. 89a."
distinction between calculating the amount of damages and fixing liability, concluding that for purposes of the latter the parties’ perceptions of the probabilities of prevailing in the litigation were more important than the actual probabilities.

The merits of the patent litigation may be crucial in an action for damages but we are here concerned only with legal liability, and we focus on the state of the world as it was perceived by the parties at the time that they entered into the settlement agreement, when they could not be sure how the litigation would turn out.”

Additionally, the Commission held that inquiries into the merits of the patent case would not be conclusive and that the uncertainties that firms would hold about patent litigation would chill settlements if their legality under antitrust law depended on an ex-post assessment of the settled patent litigation.

THE ELEVENTH CIRCUIT DECISION

Schering-Plough and Upsher-Smith appealed the Commission’s decision to the Eleventh Circuit Court of Appeals, and on March 8, 2005, that court handed down an opinion setting aside the Commission’s decision and vacating the Commission’s order.”

Reverse Payments

The appeals court rejected both the Commission’s factual findings concerning the particular settlements in this case and its reasoning on the subject of reverse payments in general.”

With respect to the Upsher-Smith settlement, the court agreed with the ALJ that the evidence supported the conclusion that the $60 million was a bona fide payment for the licenses.” With respect to the ESI settlement, the court found that the factual record was “far less developed” than

"FTC (2003b), pp. 83a-84a.
"’The court reviewed the FTC’s findings of fact and economic conclusions under the “substantial evidence” standard. According to the court’s description of that standard, “The FTC’s findings of fact, if supported by evidence, shall be conclusive” (Schering-Plough v. FTC, p. 1062, citing 15 U.S.C. § 45(c)). The court went on to explain that, “We may, however, examine the FTC[‘s] findings more closely where they differ from those of the ALJ. ‘Substantial evidence is more than a mere scintilla,’ and we require ‘such relevant evidence as a reasonable mind might accept as adequate to support a conclusion’” (Schering-Plough v. FTC, p. 1062 with internal citations omitted).
"Schering-Plough v. FTC, p. 1070.
that pertaining to the Upsher settlement, that the patent litigation was certain to be "a bitter and prolonged process," and that the settlement was "within the patent’s exclusionary power."

The appeals court did not attempt to weigh the relative merits of the economic analyses of incentives put forward by the expert witnesses for the two sides. Instead, the court flatly rejected the view that any economic analysis of incentives was an adequate basis to conclude that settlement agreements were unreasonable restraints of trade.

The Commission did not expressly adopt [CC’s expert witness]'s theories, but his rationale and the Commission's conclusions became one and the same. The Commission is quite comfortable with assenting to [the witnesses]' rather amorphous "incentive" theory despite its lack of empirical foundation. Unfortunately, [the witnesses]'s so-called incentives do not rise to the level of legal conclusions. We understand that certain incentives may rank high in these transactions, but it also true that the possibility of an outside impetus often lays dormant. The simple presence of economic motive weighs little on the scale of probative value.”

Likewise, the appeals court rejected the Commission's own reasoning on the subject.

Although it claimed to apply a rule of reason analysis,... the Commission pointedly states that it logically concluded that "quid pro quo for the payment was an agreement by the generic to defer entry date beyond the date that represents an otherwise reasonable litigation compromise.” We are not sure where this “logic” derives from, particularly given our holding in Valley Drug."

In fact, the appeals court viewed reverse payments as natural consequences of the Hatch-Waxman apparatus. The court pointed out that prior to Hatch-Waxman a generic manufacturer would need to enter the market and begin producing its product in order to challenge an incumbent's patent. Doing so would be a risky proposition for the generic because the damages to which it would be subject are the lost monopoly profits of the patentee, which are likely to be greater than any profits that the generic entrant earns.

Under Hatch-Waxman, by contrast, a challenge to a patent may be initiated simply by filing an ANDA with a Paragraph IV certification. The court took the view that this change altered the balance of bargaining power in favor of patent challengers, so that it was more likely that settlements would involve payments to the generic challenger.” The court did not make clear what implications this argument has for the question of whether or not

---

*Schering-Plough v. FTC, pp. 1071-1072.
*Schering-Plough v. FTC, p. 1069.
*Schering-Plough v. FTC, p. 1073.
settlements with reverse payments are anticompetitive except to observe that in this setting treating all settlements with reverse payments as antitrust violations would discourage settlements.”

**Per se and the Rule of Reason**

In view of its analysis of reverse payments, it is hardly surprising that the court declined to find reverse payments per se illegal. The court also declined to apply the rule-of-reason analysis, concluding, “We think that neither the rule of reason nor the per se analysis is appropriate in this context.” The court concluded that neither analysis was appropriate because the objective of both is to determine if the challenged conduct has an anticompetitive effect on the market. The court concluded that an anticompetitive effect is necessarily present because patents are, by their nature, exclusive and preferred to focus on whether or not the effect of the challenged conduct went beyond the reach of the patent. Specifically, the court concluded that an analysis of antitrust liability should include “(1) the scope of the exclusionary potential of the patent; (2) the extent to which the agreements exceed that scope; and (3) the resulting anticompetitive effects.”

Applying this standard, the court took notice of the facts that there had been no allegations that either the 743 patent was invalid or that Schering’s patent suits had been shams. On the presumption that the patent is valid the court concluded that the proper analysis was “whether there is substantial evidence to support the Commission’s conclusion that the challenged agreements restrict competition beyond the exclusionary effects of the 743 patent.” For the reasons described above, the court concluded that there was not.

**Evaluating the Underlying Patent**

Echoing the logic of Schering’s expert witness, the appeals court concluded that, because reverse payments are not intrinsically anticompetitive, it was necessary to evaluate the underlying patent litigation.

Simply because a brand-name pharmaceutical company holding a patent paid its generic competitor money can not be the sole basis for a violation of antitrust law. This alone underscores the need to evaluate the strength of the patent.

---

*Schering-Plough v. FTC,* p. 1074.

*Schering-Plough v. FTC,* p. 1075.

*Schering-Plough v. FTC,* p. 1065.

*Schering-Plough v. FTC,* p. 1066, citing *Valley Drug*, 344 F.3d at 1312.

*Schering-Plough v. FTC,* p. 1068.

*Schering-Plough v. FTC,* p. 1075-1076.

*Schering-Plough v. FTC,* p. 1076.
CERTIORARI

In August 2005 the FTC asked the Supreme Court to consider hearing the case (i.e., to grant certiorari). Revealing a disagreement between the two agencies, the Department of Justice (DOJ) filed a brief in May 2006 asking the Supreme Court not to take the case.

It appears from its petition that the most urgent issue to the Commission in pressing its Supreme Court appeal was its concern over how the appeals court language concerning the "scope of the exclusionary potential of the patent" is to be understood. There are at least three different ways in which that phrase could be understood. First, the exclusionary scope of the patent could be determined by the actual probabilities of the parties prevailing in the patent litigation. From this perspective, a settlement with a negotiated entry date exceeds the exclusionary scope of the patent if the agreed-upon entry date is later than the expected date of entry under litigation. By the second interpretation of the exclusionary scope of the patent, a settlement exceeds the exclusionary potential of the patent if the parties believe they are moving back the expected entry date. A third interpretation of the exclusionary scope of the patent would simply be its life under the assumption that the patent would be upheld in the patent litigation. By this standard, any settlement that included negotiated entry during the life of the patent would not exceed the exclusionary scope of the patent.

The Commission's certiorari petition articulates a concern that the appeals court decision would be interpreted as applying the third interpretation of the scope of the patent.

[The Appeals Court] failed to appreciate that parties settling litigation deal in uncertainties . . . The standard the court set down gives patentees free rein to 'buy off' potential competitors. The sweeping nature of the court of appeals' rule derives from its approach to assessing the "exclusionary potential of the patent." The court based its reasoning upon the statutory presumption of patent validity, and upon a demonstrably incorrect extension of that presumption to the patent infringement issues most relevant here, and ruled that the "exclusionary power" of the patent at issue here encompassed a right to exclude both Upsher and ESI from the market "until they proved either that the '743 patent was invalid or that their products . . . did not infringe Schering's patent."

"FTC (2005b)",
The Commission makes it clear that its preferred interpretation of "the exclusionary scope of the patent" is the second one above, to assess the exclusionary scope of the patent by the parties' perceptions and to discern the perceptions as inferences from the parties' actions.

The Department of Justice appeared to read the appeals court decision less drastically. For example, in its amicus curiae the DOJ observed that the appeals court found the ESI settlement within the patent's exclusionary power only after observing that the FTC had not rebutted testimony from a Schering expert witness that Schering would have won the underlying patent case. Thus, this decision is consistent with either the first or the third interpretation of the scope of the patent. More generally, the DOJ summarized its reading of the appeals court decision by asserting that it did not foreclose reliance on an ex ante view of the strength of the infringement claim in determining whether the settlement, including any reverse payment, delayed entry beyond the exclusionary potential of the patent.

In its request for a Supreme Court hearing, the Commission identified reverse payments as the central issue and sought to pose the question in a way that assumed that they are anticompetitive. In dramatic contrast, the DOJ described the principal issue for Supreme Court review in a fashion that left open the characterization of the impact of reverse payments. Here is their language for the sake of the comparison:

**FTC Question Presented**

Whether an agreement between a pharmaceutical patent holder and a would-be generic competitor, in which the patent holder makes a substantial payment to the challenger for the purpose of delaying the challenger's entry into the market, is an unreasonable restraint of trade.

**DOJ Question Presented**

Whether the antitrust laws prohibit a brand name drug patent holder and a prospective generic competitor from settling patent infringement litigation by agreeing that the generic manufacturer will not enter the market before a future date within the term of the patent and that the patent holder will make a substantial payment to the generic manufacturer.

In June of 2006 the Supreme Court declined to hear the case. The Court did not explain the reasons for its decision to deny certiorari, and the case ended.

REFLECTIONS

From the beginning, the FTC—either through its Complaint Counsel or, later through the Commission itself—maintained two positions. First, reverse payments in patent litigation settlements are intrinsically anticompetitive; and second, it is neither necessary nor possible to inquire into the strength of the underlying patent in order to determine whether or not a settlement is anticompetitive. The two positions are related. The first can provide a justification for the second. The Commission’s position, which relied on a conclusion from a particular economic theory embraced by its expert witness, was attacked on the economic grounds that the theory was wrong and on the legal grounds that the economic theory is insufficient for the rule of law that the Commission advocated.

By the end of the case, the Commission was arguably in a worse position than it had been in when the case began. The Commission found itself attempting to preclude the imposition of a rule from the circuit court that, at least as the Commission saw it, would permit any settlement to patent litigation that allowed entry within the nominal life of the patent. If the Commission believed that this state of affairs had been reached because it had persisted in pursuing its two positions together, it might have reason to regret ever having taken up the case in the first place.

At this juncture, it is of greatest importance to set aside for the moment the question of whether or not one believes that the Commission’s or the Respondent’s arguments were the stronger in this case and to address the long-run issue of the best policy toward agreements that settle patent litigation. While it is not our ambition to resolve this issue here, the history of the FTC’s litigation against Schering-Plough seems to have illuminated a few lessons that bear on any such resolution:

(1) The parties to a patent dispute will generally have incentives to settle in order to avoid costly and risky litigation, and these incentives are more aligned than not with social benefits. (2) It does follow from economic logic that the parties to a patent dispute both could benefit from a settlement that significantly protected the incumbent patent holder from the chance of competition from the accused infringer and provided a net payment in return from the patent holder to the accused infringer. As such, the benefits to the parties would come at the expense of competition and consumer welfare. (3) It also follows from economic logic that the opportunity to employ reverse payments may be necessary for socially beneficial and pro-competitive settlements to be reached, due to such common situations as asymmetric information, excess optimism, and differential cash needs between the parties to the patent dispute. (4) It may very well be the case that it is not feasible for an antitrust inquiry to reach reliable judgments about the likelihoods of the litigation outcomes of a patent dispute.

These four lessons have implications for policy. They seem to imply that a per se rule against reverse payments would, at this point in our inexperience,
be unduly restrictive of settlements and hence unwise. They imply that a safe harbor for agreements that settle patent disputes would foster anticompetitive outcomes that benefit the parties to the patent disputes but that harm consumers. They imply that we cannot rely on a quantitative comparison between the agreed date of competitive entry and an assessment of the odds of the possible outcomes of the patent litigation.

The remaining implication is that more experience by antitrust agencies and courts, along with more economic research, is needed to formulate indirect tests for whether a settlement is likely to benefit or harm competition. While we may not be able to measure the probabilistically expected date of entry under litigation, we may nevertheless be able to assess whether the date of entry under the settlement is likely to come before or after that expected date on the basis of the character of the settlement negotiations and their context, as well as on the basis of the size and compensations for any reverse payments. The resulting challenge clearly needs carefully crafted, articulated, and tested "second best" solutions. If the legacy of the FTC v. Schering-Plough cases is the motivation for aware and talented economists and lawyers to tackle this problem, then that could be the best outcome from the whole matter.

REFERENCES


Page 275
INTRODUCTION

In 1997, LePage’s Inc., a small manufacturer of home and office tape, filed a lawsuit in the U.S. District Court for the Eastern District of Pennsylvania, alleging that 3M had “monopolized or attempted to monopolize the invisible and transparent home and office tape market by squeezing out its small competitors, including plaintiff LePage’s through, among other things, monopoly leveraging, exclusive dealing, bundled rebates, full line forcing and other predatory conduct.” In 1999, after a nine-week trial, a Philadelphia jury found in favor of LePage’s and assessed damages that, after trebling, exceeded $68 million.

The decision was appealed by 3M to the U.S. Court of Appeals for the Third Circuit. In 2002, a three-judge court of appeals panel reversed the district court decision on a divided vote and instructed the district court to enter judgment in favor of 3M. However, the initial court of appeals opinion was subsequently vacated and the appeal reheard by the entire court of appeals (an en banc hearing), which in March 2003 affirmed the original district court decision. 3M petitioned for a hearing before the U.S. Supreme Court but was denied.
The decision in LePage's v. 3M was highly controversial. Some observers applauded the court’s efforts to rein in what are perceived as targeted efforts to undermine competition. Critics, however, argued that the decision does not provide a coherent basis for distinguishing possible predatory or exclusionary conduct from vigorous competition and efficiency-enhancing business activities and that the resulting ambiguity threatens to chill a wide range of pro-competitive activities.

This chapter describes the dispute between LePage’s and 3M, the associated court of appeals decision, and the significance and implications of the case. It first provides the basic factual background necessary and then describes the various arguments made by LePage’s and 3M and the evidence presented by the parties to support their claims. The chapter next summarizes the district court decision and the subsequent opinions offered by the court of appeals. It concludes with some observations regarding the implications of these decisions for the evaluation of unilateral firm conduct.

BACKGROUND

3M is a large, diversified company that produces a wide range of products used in industrial, commercial, and consumer applications. 3M consumer products are sold in a variety of retail outlets, including mass merchants, drug stores, grocery stores, convenience stores, and office supply stores. The 3M consumer products are organized into a number of categories, including stationery products (e.g., home and office tape, repositionable notes), packaging products, home care products, leisure products, and audiovisual products. Several of the 3M consumer products are very successful. For example, 3M is (by far) the largest supplier of home and office tape in the United States; indeed, during the early 1990s (the time period when the dispute between 3M and LePage’s occurred) 3M accounted for over 90 percent of all home and office tape sales in the United States.

3M sells two brands of transparent home and office tape. The Scotch brand is promoted for home use and sold primarily through consumer retail outlets. The Highland brand is promoted for business use and sold primarily through office supply stores and, to a lesser extent, through mail-order outlets. Historically, 3M did not generally participate as a supplier of private label home and office tape, as it preferred to focus its efforts on its branded products. However, as the demand for private label products grew (particularly with the growth of the large office product “superstores” such as Staples and Office Depot), 3M reconsidered this decision and began to

Private label products are manufactured by one company (in this case 3M or LePage’s) but offered for sale under another company’s brand name (usually a store brand that is unique to an individual retailer). Prior to 1992, 3M had only two private label home and office tape customers.
offer private label home and office tape in 1992. By 1997, 3M accounted for about one-third of all private label home and office tape sold in the United States through sales to a few dozen very large customers.

In 1993-1994, 3M introduced an experimental marketing program designed to encourage 3M customers to increase their purchases of 3M products. Under the "executive growth fund" (EGF) program, 3M negotiated annual growth-in-purchase targets for each of the six 3M consumer product divisions with each of the 13-15 largest customers. Each was eligible to receive a rebate on its total annual 3M purchases scaled according to the number of 3M consumer product divisions for which the growth targets were achieved.

In 1995 3M replaced the EGF with the "partnership growth fund" (PGF). The PGF was available to all purchasers of 3M consumer products and involved a common menu of rebates (ranging from 0.5 to 2 percent) conditioned on the growth in the customer's annual purchases from 3M overall and the number of consumer product divisions for which customers increased their purchases relative to the prior year. Customers who failed to achieve an overall growth rate of 12 percent or to increase their purchases from at least two 3M consumer product divisions would receive no year-end discount.

LePage's began to sell home and office tape in the United States around 1980, both under the LePage's brand name and under private label. LePage's achieved some success, particularly as a supplier of private label products. During the early 1990s, it accounted for almost 90 percent of all private label home and office tape sold in the United States and for the vast majority of all home and office tape sold in the United States by firms other than 3M.

After 3M began to sell private label home and office tape, a number of LePage's customers switched suppliers and began to purchase private label home and office tape from 3M. LePage's share of private label tape sales declined significantly, and the company subsequently filed its antitrust suit claiming that 3M had employed a number of anticompetitive tactics designed to eliminate competition from LePage's and other alternative suppliers of home and office tape in order to protect 3M's monopoly position.

3M's decision to begin offering private label home and office tape more generally to its customers was facilitated by the recent opening of a new manufacturing facility that reduced 3M's production cost for private label home and office tape.

The EGF and PGF were not the only two 3M marketing programs that linked the prices of its products to the range and volume of 3M products purchased by its customers. They were, however, typical of the others and were the focus of LePage's complaint against 3M.

There were one other small U.S. supplier (Tesa Tuck, Inc.) and a number of foreign suppliers. None of these firms had substantial sales in the United States, and Tesa stopped supplying home and office tape before the trial began.

There were four separate allegations: (1) unlawful maintenance of monopoly power, (2) unlawful attempted maintenance of monopoly power, (3) unlawful exclusive dealing, and (4) unlawful restraint of trade. See LePage's Second Amended Complaint.
LePage’s claimed that 3M’s expansion into private label home and office tape was anticompetitive because it was motivated by 3M’s desire to control and destroy the low-price private label segment of the home and office tape market. According to LePage’s complaint, 3M intended to accomplish this goal by targeting LePage’s’ most important customers for conversion to 3M while limiting the availability of private label home and office tape elsewhere in the market.

THE ARGUMENTS

As in all monopolization cases, LePage’s needed to establish two component propositions. First, LePage’s needed to demonstrate that 3M possessed, or was in danger of acquiring, monopoly power in a relevant market. Second, LePage’s needed to show that 3M’s monopoly power was, or would be, willfully acquired or maintained through anticompetitive means. That is, LePage’s needed to show that 3M had obtained, or threatened to obtain, monopoly power through some form of predatory or exclusionary conduct and not merely through its superior products or business acumen. Most of the controversy in the case focused on the second proposition; that is, the characterization and effects of the 3M marketing strategies. Nevertheless, it is instructive to describe at least briefly the parties’ positions with regard to market definition and market power, as these factors establish the environment in which the conduct in question must be evaluated.

The Relevant Market

For the purposes of analyzing antitrust claims, relevant markets are defined by the products that are close substitutes for the products in question (in this case, the home and office tape produced by 3M and LePage’s) and the geographic area where those products are bought and sold. LePage’s argued that the relevant product market for the purpose of analyzing its antitrust claims against 3M was simply the manufacture and sale of transparent tape for home and office use. LePage’s went on to argue that the only relevant sellers in this market were domestic producers and that thus the relevant geographic area was the United States. LePage’s claimed that foreign suppliers could not sell tape effectively in the United States because they sold polypropylene tape, which has somewhat different characteristics than the acetate tape sold by 3M and LePage’s (and the cellophane tape that was also sold by LePage’s). In support of this proposition, it cited to the limited sales made by foreign manufacturers in the United States.

While 3M did not dispute the general claim that transparent tape for home and office use in the United States was a relevant antitrust market, it did contest LePage’s argument concerning foreign suppliers. 3M claimed
that product differences did not prevent foreign manufacturers from selling home and office tape in the United States or from competing effectively with 3M and LePage’s. 3M further argued that foreign suppliers were particularly attractive alternatives in the discount segment of the market, where price was an important element of customer choice and private label products predominated. 3M cited to evidence that both 3M and LePage’s had offered price discounts in response to competition from foreign suppliers, and 3M identified a U.S. customer who had recently added a line of private label tape manufactured in Asia. 3M also argued that the discounts offered in response to the threat of foreign competition demonstrated that foreign firms could constrain prices (and prevent incumbent suppliers from increasing prices) even without capturing a significant share of sales in the United States.

Monopoly Power

Monopoly power is commonly defined as the power to control price or to exclude rivals from a relevant antitrust market. LePage’s argued that 3M possessed monopoly power in the relevant market for home and office tape. LePage’s supported this claim by pointing to a number of frequently cited indicia of market power, including 3M’s large share of total sales in the relevant market and its high reported return on invested capital. LePage’s also identified a number of potentially significant impediments to entry. These included the considerable time and cost required to develop a viable brand name, the existence of significant economies of scale in production that limited the viability of small-scale entry, the high capital investment requirements that excluded many possible entrants, and the difficulty of establishing effective distribution channels. All of these were said to insulate 3M from the threat of competition from new entrants.

3M disputed these various claims. 3M pointed out that high market share does not necessarily imply market power, since share may simply reflect high quality, good service, and low prices. Further, that share was continually at risk from competition both from LePage’s and from a number of foreign firms that were capable of and in some cases already were supplying home and office tape in the United States. 3M also pointed out that after six years of head-to-head competition with 3M, LePage’s still accounted for the vast majority (over 75 percent) of all private label home and office tape sales in the United States. With regard to 3M’s high reported return on capital, 3M argued that accounting profitability, like market share, cannot be relied on as proof that a firm has substantial monopoly power, since the value of many capital assets (e.g., brand name, intellectual property, land, factories, and equipment) are frequently underreported in financial statements, and that reported returns may thus overstate a company’s true return on invested capital.

3M further argued that it was constrained by the power of its largest customers, who accounted for a significant portion of total 3M sales. These customers were said to be sophisticated purchasers who were willing and
able to switch suppliers in the event that 3M attempted to exercise monop-
oly power. 3M highlighted a number of instances in which 3M customers
extracted discounts by threatening to purchase home and office tape from
LePage’s and other suppliers, including firms that did not currently supply
home and office tape in the United States. 3M suggested that these cus-
tomers were capable of sponsoring entry by new suppliers or vertically
integrating into home and office tape production themselves.

3M also disputed LePage’s claim that entry into the United States by a
new supplier of home and office tape was unlikely due to the high cost and
substantial time required to effect entry. 3M noted that LePage’s claim that
a new entrant would have to establish a strong brand identity in order to
compete effectively in the United States was inconsistent with the fact that
LePage’s own business was to supply private label and secondary brand
products and that its complaints centered on 3M’s supply of private label
rather than branded home and office tape. 3M also argued that LePage’s
claim regarding economies of scale was similarly inconsistent with
LePage’s own success, since it never achieved a particularly large share of
home and office tape sales in the United States. 3M disputed the claim that
capital investment requirements constituted a substantial barrier to entry by
identifying several large customers (e.g., Staples) whom it believed were
able and willing to bear the necessary capital investment requirement indi-
vidually. 3M also identified several firms that sold significant quantities of
home and office tape outside the United States; for these firms sales in the
United States would be incremental and thus not subject to the same scale
disadvantages of de novo entry.

Willful Maintenance of Monopoly Power
LePage’s central antitrust claim was that 3M had maintained its monopoly
position in the relevant market by improperly preventing its smaller rivals,
most significantly LePage’s, from competing for sales through a variety of
unfair competitive activities. This claim was supported by three types of
evidence. First, LePage’s offered evidence to suggest that 3M had the
motive and intent to maintain the high prices of its branded home and office
tape products by restricting consumer access to private label or secondary
brands. Second, LePage’s offered a theoretical framework (the theory of
monopoly preemption) that explained how a firm with monopoly power
might attempt to accomplish such a goal. Third, LePage’s attempted to
show that 3M’s conduct, and particularly its various bundled rebate pro-
grams, had the practical effect of accomplishing this anticompetitive objec-
tive. We discuss each of these in turn.

3M’s Motive and Intent
The first element of LePage’s case against 3M focused on 3M’s motives
for entering as a supplier of private label home and office tape in 1992.
During the 1980s, the prospects for private label home and office tape sales improved significantly due to the growth of mass merchandisers such as Wal-Mart and Kmart, office superstores such as Staples and Office Depot, and membership clubs such as Sam’s Club and Costco. Over this period, retailers employing these formats developed the sales volume and customer acceptance necessary to sponsor their own brands of home and office tape. LePage’s offered evidence at trial to demonstrate that 3M recognized that the growth of these retail formats enhanced the potential for private label competition.

LePage’s identified three ways in which the growth in demand for private label home and office tape threatened 3M’s business: (1) by eroding 3M’s share of home and office tape sales; (2) by putting downward pressure on the prices of 3M’s branded home and office tape products; and (3) by providing an opportunity for LePage’s to develop into a stronger and more formidable competitor overall. Evidence was presented to show that 3M expected increasing competition, particularly competition involving private label products, to force reductions in the prices of premium branded products such as Scotch brand home and office tape and Post-it brand repositionable notes. Evidence was also presented to show that 3M sales of branded products in other product categories (e.g., mailing tape) had been previously eroded by competition from secondary brand and private label products. Based on this evidence, LePage’s concluded that 3M knew it had much to lose if competition from private label and secondary brand home and office tape continued to develop and had a strong incentive to reduce or eliminate such competition.

LePage’s claimed that 3M’s approach to becoming a supplier of private label home and office tape demonstrated that its motive was to protect the monopoly profits it obtained from sales of its branded products and not to compete legitimately in the private label segment. Evidence was presented that 3M intended to sell private label home and office tape only on a limited basis and only to customers who were already purchasing private label home and office tape from a competitor or seriously considering a private label program. LePage’s argued that 3M hoped ultimately to reduce (“kill”) the penetration of private label products by switching customers back to its branded products after it became established as the customers’ supplier of both branded and private label home and office tape.

LePage’s also identified a number of instances in which 3M encouraged its branded home and office tape customers to increase the retail prices of those products. LePage’s argued that these efforts were directed at retailers who were starting to consider offering less expensive private label home and office tape. Evidence was presented to show that 3M believed that if it could convince those retailers to increase the retail prices of Scotch brand home and office tape products, it would thereby reduce the interest of those retailers in carrying or increasing their promotion of private label products.
Thus, LePage's argued that 3M's entry as a supplier of private label home and office tape was not the kind of competition that the antitrust laws were intended to support or protect. To the contrary, LePage's argued that 3M began selling private label home and office tape only in order to maintain its monopoly prices and restrict output and customer choice by preventing the market from evolving toward lower-priced private label products.

3M responded that its decision to manufacture private label home and office tape in addition to its existing branded products was a legitimate competitive strategy that expanded customer choice, increased quality, and decreased prices. 3M presented evidence to show that some customers selected 3M over LePage's because of its attractive prices, service, and quality, while others used 3M's offers to supply private label home and office tape, or its mere presence as an alternative supplier of private label products, as a tool for negotiating lower prices from LePage's.

3M further explained that its desire to promote its branded home and office tape products more aggressively than its private label products reflected its legitimate business interests. 3M recognized that additional private label sales might displace its more profitable branded product sales to some extent. Given its interest in selling branded home and office tape, 3M emphasized its branded products and focused its efforts to sell private label home and office tape on customers who were not interested in its branded products or had demonstrated a predisposition to purchase private label products.

LePage's claimed that 3M specifically targeted LePage's customers, with the intention to drive LePage's from the marketplace. In response, 3M explained that when it began to manufacture private label home and office tape, it naturally approached customers who had previously demonstrated a preference for private label products. The obvious place to seek incremental private label home and office tape sales was from customers who had previously purchased private label home and office tape from LePage's or other suppliers. 3M characterized this as an obvious and legitimate business strategy. Moreover, 3M claimed that it did not specifically target LePage's customers. Evidence was presented to show that 3M supplied private label home and office tape to a number of customers who were not previously supplied by LePage's. Although some of 3M's initial private label customers had previously purchased from LePage's, twice as many had purchased from Tesa or a foreign supplier or had not previously purchased private label home and office tape products at all.

The Primary Theory of Harm: Monopoly Preemption

LePage's theory of harm was based on the economic theory of preemption. While anticompetitive preemption may occur in a number of ways, the theory alleged in this case was that 3M used its control over a
scarce resource (critical retail distribution channels) to prevent others from competing, thus restricting output and obtaining monopoly profits. An incumbent monopolist may be in a unique position to bid such a scarce resource away from its rivals for the following reason: The most that a potential entrant or fringe competitor is willing to pay for the resources required to effect entry or expansion is the profit that can be earned if the entering or expanding firm is successful. The scarce resource is generally worth more to the incumbent monopolist because such a firm stands to gain the entire monopoly profit if it can prevent entry or expansion, in contrast to the fraction of profits that the potential entrant can hope to obtain. The primary factors determining the extent of profitable preemption are the number of different suppliers of the scarce resource, the amount of the scarce resource held by the incumbent firm and each of the alternative suppliers, and the downstream market demand elasticity.

LePage's argued that 3M's various marketing programs preempted LePage's by giving 3M control over scarce distribution channels for its products. Specifically, 3M was alleged to penalize those retailers who might otherwise have carried Scotch brand tape plus LePage's by offering benefits to customers who substituted away from LePage's in favor of 3M's private label tape. In short, by distorting the stocking decisions of its customers, 3M secured control over the distribution channel and locked out LePage's products.

3M denied that contention. It noted that customers were not prohibited from purchasing any product from any supplier, including the private label home and office tape supplied by LePage's. It further argued that the actual effect of its marketing programs was to lower prices and thereby encourage customers to increase their purchases of 3M products. 3M argued that displacement should not be confused with preemption, since competitors are routinely displaced whenever a new firm offers better value to contested customers. 3M claimed that that was precisely what it had done wherever it displaced LePage's as the supplier to a particular customer. 3M argued that if such price-based competition is to be evaluated under the antitrust laws, it should be analyzed as pricing conduct—predation—rather than under some misplaced theory of preemption.

3M further argued that the theory of preemption was inappropriate in this instance because the resource in question—access to retail consumers and business customers—is not actually scarce. A large number of independent retailers provide access to consumers and business purchasers. Moreover, individual retailers can vary the resources (e.g., shelf space) devoted to home and office tape and can purchase from more than one supplier if they choose. LePage's in turn argued that not all retail outlets were equally attractive and that 3M preempted LePage's and other competitors from the retail outlets that were most critical to the competitive success of these smaller rivals.
Evaluation of the Contested 3M Marketing Strategies

In support of its theory, LePage’s asserted that 3M employed a variety of tactics to gain exclusive supplier status at critical retail distributors or to induce them to reduce their purchases of private label and secondary brand home and office tape from other manufacturers. These tactics included offering bundled rebates, growth rebates, brand mix rebates, and other financial incentives. LePage’s argued that all of 3M’s marketing programs shared a common, exclusionary feature: They explicitly or implicitly linked discounts on products where 3M possessed market power (e.g., Scotch brand tape and Post-it brand notes) to the customer’s decisions regarding the purchase of private label home and office tape. LePage’s further argued that 3M offered no credible nonexclusionary rationale for its programs, or at least no rationale that could not be realized without far less harm to competition and consumers.

3M Bundled Rebate Programs

The various 3M rebate programs at issue shared the same basic structure: In each instance the customer was able to obtain discounts (rebates) on a range of products by achieving sales or growth objectives overall and (in some cases) with regard to smaller subsets of the full product set. The plans thus potentially created circumstances in which the customer’s decision with regard to the purchase of a single product or product line could affect the discount obtained not only on that particular product but also on a range of other products. LePage’s argued that the effect of such programs was to levy a very substantial lump sum tax on the activities of smaller rivals such as LePage’s.

LePage’s logic was as follows: Customers had no practical choice but to buy Scotch (and Post-it) brand products (the products for which LePage’s claimed 3M enjoyed substantial monopoly power) because there were no competing premium brands. However, many customers purchased (or were willing to consider purchasing) a portion of their home and office tape requirements (e.g., secondary brand or private label home and office tape) from other vendors such as LePage’s. The 3M rebate programs interfered with these purchasing decisions by imposing a potentially large lump sum tax on customers who patronized vendors other than 3M. The effect of the programs was to create incentives for customers to displace LePage’s secondary brand or private label home and office tape with branded or private label 3M products or to resist shifting a portion of their home and

In several instances LePage’s suggested that 3M also had monopoly power in the supply of repositionable notes due to the strength of the 3M Post-it brand products. LePage’s did not expressly define a relevant market for the supply of repositionable notes or establish that 3M had monopoly power in such a market.
office tape requirements from 3M to LePage's in order to maximize their chances of meeting the overall growth targets and narrower divisional growth targets required to obtain rebates on their purchases of 3M products.

LePage's offered evidence showing that 3M sometimes recommended that customers replace LePage's with 3M as the supplier of their private label home and office tape products or discontinue sales of LePage's private label home and office tape in order to meet their growth targets for 3M's consumer stationery division. Some customers indicated that one of their objectives was to manage their procurement in order to qualify for the maximum possible 3M rebates. Given the structure of the 3M rebate programs, failing to meet the volume growth target for any one 3M consumer product division could reduce the amount of the rebate received on all 3M products or eliminate the rebate altogether.

To illustrate, LePage's provided the following example: Suppose a customer was initially purchasing $1 million worth of 3M products from each of five 3M consumer product divisions. Suppose also that transparent home and office tape accounted for $500,000, or one-half of the customer's purchases from one of the divisions (the consumer stationery division)." Finally, suppose that 3M offered a 0.75 percent rebate on all of the customer's purchases if the customer continued to purchase at least $1 million worth of 3M products from each of the 3M consumer product divisions. That is, 3M set the "growth targets" at exactly the customer's existing purchase levels. Now suppose the customer was considering an offer (by LePage's) to replace 20 percent of its Scotch brand home and office tape purchases with LePage's private label tape products, and LePage's had offered to supply the private label products for 40 percent less than the price that 3M was charging for similar Scotch brand products.

On its face, it would appear that LePage's was offering the customer a potentially attractive opportunity to obtain private label home and office tape that it could sell at a significant discount compared with the price of the branded 3M products while simultaneously increasing its retail profit margin. The net cost difference of $39,250 ($40,000 less the 0.75 percent rebate 3M was offering directly on its branded home and office tape products through the rebate program) could be used to reduce consumer prices, increase the retailer's profit, or some combination of the two. From the customer's perspective, however, the value of the LePage's offer would be far less, due entirely to the structure of the 3M rebate program. If the customer accepted the offer from LePage's, it would fail to meet its volume growth target for 3M's consumer stationery division and thus forfeit the

"Transparent home and office tape accounted for approximately 50 percent of 3M's total sales of the consumer stationery products, one of the six 3M consumer product divisions.

"It is assumed in this example that any quality differences between the branded and private label products are small in relation to the magnitude of the price difference between the branded and private label products.

Page 286
anticipated rebate of $37,500 (0.75 percent times $5 million). Thus, the effective net cost difference would be only $2,500, or 2.5 percent. This might not be sufficient to overcome the (perceived) difference in the quality of the branded and private label products.

LePage’s argued that LePage’s and other sellers of private label home and office tape could not compete against the 3M bundled rebate programs. LePage’s reported that its gross margin on home and office tape products was about 33 percent of sales, or about $20,000 on a sale of $60,000 worth of merchandise, and its net profits (after deducting its fixed overhead costs) were only about 8 percent of sales, or $4,800 on the same $60,000 sale. Thus, the lump sum payment or additional price concession that LePage’s would have to offer the customer in order to compensate for the forgone rebate would be almost twice as large as LePage’s gross profit on the sale, and almost eight times its net profit.

A similar calculus was applied to situations in which LePage’s (or another manufacturer) was already supplying a portion of the customer’s home and office tape requirements. Suppose the same customer had previously displaced $100,000 worth of branded 3M home and office tape by purchasing $60,000 worth of private label product from LePage’s (or another manufacturer), and 3M offered the same 0.75 percent rebate on all of the customer’s purchases from 3M if the customer increased its purchases of 3M products from the existing level of $4.9 million to $5 million (an increase of about 2 percent). In this case, the customer might feel compelled to discontinue purchases of private label home and office tape and increase its purchases of the 3M branded products in order to attain the 0.75 percent discount. To retain the customer, LePage’s would have to offer the customer a lump sum payment or additional discounts worth at least $37,500. This amount would again be almost twice as large as LePage’s gross profit on the sale and almost eight times its net profit.

LePage’s argued that the example demonstrated that the 3M rebate plans were anticompetitive because their structure prevented an efficient single-line competitor from profitably competing for sales that it could otherwise obtain. Even a small percentage rebate applied to the large base of all 3M sales could make it uneconomical for the customer to do business with LePage’s or another narrow-line supplier if the customer felt that doing so would place its 3M rebate at risk. LePage’s claimed that 3M attempted to set the rebate targets at levels that would influence customer behavior in precisely this way and reasoned that 3M had no interest in setting rebate targets that were either too easy or impossible for the customer to meet because then the programs would not affect customer procurement decisions in the desired way.12

In some instances (e.g., the PGF program), the relevant sales or growth targets were common across all 3M customers. Although LePage’s acknowledged that individually customized sales or growth targets provide more precise tools for exclusion, it maintained that the 3M programs with
3M argued that its various rebate programs were nothing more than ordinary price reductions. Although the various 3M rebate programs differed in their details, they shared a common characteristic: They all reduced the prices that 3M charged to customers who purchased specified volumes of 3M products or increased their purchases over time. In other words, the rebates were simply price discounts, eligibility for which was based on the customer’s volume and growth characteristics.

3M produced several examples to show different ways by which customers might calculate the value of the discounts depending on the purpose of the calculation. For example, in order to compare the offers made by competing suppliers, customers could merely calculate the prices of all 3M products net of the expected rebates. Thus, the customers in the LePage’s example could determine the true net prices of the products they purchased from 3M by merely deducting the value of the rebate they expected to receive (0.75 percent) from the price of each product. These net prices could then be compared with the prices offered by competing suppliers (similarly net of any applicable rebates) to determine which supplier(s) made the best offer(s).

For some purposes, it might be appropriate to allocate the rebates to products other than those to which they naturally applied. For example, the customer in the first LePage’s example could apply the entire amount of the expected 3M discount ($37,500) to the $100,000 purchase of branded 3M home and office tape that would be displaced by the purchase of private label products in order to reflect the full value of the discount that would be forfeited if the customer reduced purchases of branded 3M products in order to purchase private label home and office tape from LePage’s.

Similarly, a firm that was considering expansion of its relationship with 3M could calculate the expected incremental cost of additional purchases net of any rebates the firm would obtain as a result, even if the rebates were applied to other products. Thus, the customer in the second LePage’s example could apply the entire amount of the expected 3M discount ($37,500) to the additional $100,000 purchase of branded 3M home and office tape that would be required to meet the volume growth target and thereby obtain the discount. The fact that the rebate would be nominally paid on other products did not prevent the customer from making the appropriate calculation. 3M asserted that each of these calculations provided a potentially legitimate conversion of the rebate payments into price.

Standardized targets were nonetheless exclusionary because they presented critical LePage’s customers with volume growth targets that made it risky or uneconomical for the customers to deal with LePage’s.

3M argued that if a customer was not sure that the 3M volume targets would be met, the amounts of the expected price reductions could be discounted to reflect the uncertainty.

Similarly, in order to evaluate true marginal product costs for the purpose of setting retail prices, unit costs could be calculated net of any applicable rebates.
3M claimed that the rebate programs lowered prices and thus induced consumers to purchase greater volumes of 3M products and that 3M expanded the sales of both its private label and branded product customers. 3M presented evidence that showed that the prices of 3M home and office tape products fell in real terms (net of inflation) for the retail consumer channel and in absolute terms for the office channel between 1993 and 1997. The general pattern of declining prices occurred despite the introduction of new, higher-priced products. The average price of existing home and office tape products fell about 6 percent over this period, while the unit sales of 3M, LePage’s, and the overall market all increased. 3M argued that competition, and price competition in particular, is fundamental to the efficient operation of a market-driven economy and that lower prices and greater purchases are typically associated with improved consumer welfare. Thus, 3M asserted, price discounts such as those implemented by the various 3M rebate programs generally should be encouraged.

One exception to this rule concerns predatory pricing. Predatory pricing occurs when prices are reduced with the specific intent of eliminating competition and with the expectation that the investment in low prices will be recouped by charging higher prices after the competition is illuminated. It is widely accepted that antitrust policy must be carefully administered to promote competitive behavior and to avoid deterring aggressive price competition. Thus, the Supreme Court has established as a legal threshold that prices cannot be predatory unless they are set below the cost of the selling firm. 15 3M explained that if the courts were to adopt an approach that condemned prices set above the seller’s own costs, it would risk deterring the firm from competing on the basis of low but profitable prices.

In fact, LePage’s did not assert a claim of predatory pricing and did not claim that 3M’s prices were set below 3M’s production costs. Rather, LePage’s asserted that the 3M rebate offers differed from ordinary price reductions because they frequently took the form of lump sum rebates that were paid only after customers achieved the annual volume or growth targets established by 3M. LePage’s argued that 3M’s rebate programs thus effectively required smaller competitors to enter into “rigged” auctions against 3M. Smaller competitors such as LePage’s could bid for access to important distribution outlets against 3M by offering their own lump sum payments, but in order to win the business they would have to compensate the customers for the expected loss of lump sum rebates from 3M if a customer failed to meet 3M’s volume or growth targets for that customer as a result of the customer’s purchasing decision. LePage’s argued that the theory of economic preemption demonstrated that if small competitors are forced to bid against an incumbent monopolist for access to a critical

resource such as the distribution outlets at issue here, the monopolist will win and preserve its market power.

LePage’s suggested that the prices implied by the 3M bundled rebate programs should be compared with the profit margins, both gross and net, of LePage’s. LePage’s asserted that if the prices charged by 3M (net of the applicable discounts) were less than the costs of its smaller rivals, those prices and the rebate programs that implemented them should be viewed as exclusionary. The implication is that 3M should not be allowed to offer prices (net of rebates) that are lower than the costs of its smaller competitors. 3M argued that it would be inappropriate to adopt any such requirement and that competition policy should be concerned with preserving competition and not accommodating smaller, less-efficient firms.

LePage’s further claimed that the 3M end-of-year rebates were carefully designed to minimize the possibility that the rebates would be treated as ordinary cost reductions and thus passed on to consumers in the form of lower prices. 3M countered by noting that its practice of calculating and paying rebates at the end of the year was an obvious and sensible policy, given that the magnitude of the rebates was based on the total annual purchases of the participating customers. Moreover, 3M claimed that the extent to which the rebates were passed on to consumers in the form of lower retail prices is not relevant to the antitrust analysis. And in any event, 3M had no control over what its direct customers did with their rebates.

3M also argued that its size and breadth should not preclude it from competing with smaller rivals or limit the range of discount plans it can employ. The essence of LePage’s complaint was that limited-line firms such as LePage’s could not compete with multiple-line firms like 3M because they could not replicate the multiple product discounts offered by such firms. 3M disputed this conclusion and argued that LePage’s was free to compete on price, quality, and service, just like any other supplier. Customers were able to compare offers from limited-line suppliers with those offered by more diversified firms such as 3M. Although LePage’s might not be able to offer a rebate program that incorporates the range of products that 3M sells, it could compete with the prices offered by 3M in connection with such programs.

3M further argued that LePage’s offered no reasonable standard for distinguishing anticompetitive conduct from ordinary competition that enables more efficient or more aggressive firms to take sales away from their less-efficient or less-aggressive rivals. The standards suggested by LePage’s analysis, 3M argued, amounted to a claim that the programs were

"LePage’s did not actually perform these calculations systematically for the customers it claimed were foreclosed by the effects of the 3M rebate programs.

"This also raises the troubling question of how 3M could be expected to know its competitors’ costs and also, if there were more than one competitor, which competitor’s costs (the most efficient? the least efficient? some average?) should be the benchmark."
exclusionary because LePage’s could not compete effectively with 3M’s low prices. Thus, it appeared to 3M that LePage’s was merely asking the court for protection from a larger, more aggressive rival. 3M argued that such requests are inappropriate because the antitrust laws were not intended to protect inefficient competitors.

Other Allegations of Anticompetitive Conduct
LePage’s characterized 3M’s marketing programs in a number of additional ways that implied anticompetitive behavior. These included tying, exclusive dealing, and monopoly leveraging. According to LePage’s, these various nonprice strategies prevented LePage’s from competing effectively and thus enabled 3M to maintain monopoly prices. 3M denied the existence of any economic basis for these additional allegations. Since these allegations substantially overlap with the allegations involving rebates, we shall address them only briefly.¹⁸

Tying

In economics, tying occurs when a firm conditions the purchase of one product on the purchase of another. A firm that makes cups and straws, for example, may agree to sell cups only to customers who agree to purchase straws as well. But tying is commonplace—left and right shoes, engines and car bodies in autos, and so on—and generally has no adverse implications for competition. Indeed, it often confers efficiency benefits on consumers who do not wish to purchase each piece or part of a desired combination and then do the assembly themselves. However, if the selling firm has monopoly power with regard to the tying product (cups, in this example), then it is possible that the selling firm might eliminate competition in the market for the tied good (straws) by refusing to sell the tying product separately. The potential harm to competition occurs in the tied good because customers who need the tying product are prevented, or at least deterred, from purchasing the tied product from another supplier.

LePage’s claimed that 3M’s various rebate programs tied the purchase of 3M’s nonmonopoly products (e.g., private label home and office tape) to the purchase of 3M’s alleged monopoly products (e.g., Scotch brand home and office tape and Post-it brand repositionable notes). 3M argued, in response, that its rebate programs did not literally or effectively tie the purchase of any 3M products. 3M’s rebate programs did not explicitly condition the purchase of any 3M product on the purchase of any other 3M product, and 3M customers appeared free to purchase any combination of 3M products they chose and to avoid any 3M products they did not want to

¹⁸We do not address the leveraging claims here, since they raise no economic issues apart from those with respect to tying and exclusive dealing.
purchase. 3M presented evidence that many customers (including virtually all of LePage’s customers) purchased one or more of the alleged tying products from 3M and yet purchased many of the alleged tied products (e.g., private label home and office tape) solely or primarily from firms other than 3M.

LePage’s tying complaint was thus limited to the claim that the 3M rebate programs imposed a de facto tie between 3M’s “monopoly” and “nonmonopoly” products, at least for some customers. In essence, LePage’s argued that the 3M rebate programs conditioned the purchase of 3M’s monopoly products on the purchase of 3M’s nonmonopoly products by linking the rebate payments in a manner that made the purchase of LePage’s products not a viable business proposition for its potential customers. In this manner, LePage’s claimed, 3M foreclosed competition in the sale of the nonmonopoly products, including secondary brand and private label home and office tape.

3M argued that this claim was again nothing more than a complaint about 3M’s low prices, as implemented through the various rebate programs. 3M asserted that the alleged monopoly and nonmonopoly products were sold separately and on commercially viable terms and produced evidence showing that many customers purchased the alleged monopoly products but did not purchase the nonmonopoly products from 3M. LePage’s countered this evidence by alleging that only certain 3M customers were compelled to purchase nonmonopoly products (including private label tape) from 3M in order to receive discounts on other allegedly monopoly 3M products. 3M responded that this more limited allegation amounted simply to the complaint that 3M’s prices for its nonmonopoly products were in some instances set lower than LePage’s would prefer.

To illustrate this point, 3M provided the following example. Suppose firm A (3M) is a monopolist in the sale of product X (Scotch brand home and office tape) and sells product X for $2 per unit, whereas firm B (LePage’s) sells product Y (private label home and office tape) for $1 per unit. Suppose further that each customer purchases one unit of product X and one unit of product Y. Finally, suppose that firm A begins to sell product Y for $1 per unit while offering its customers a 10 percent rebate ($0.20) on each unit of product X they purchase if the customers’ total purchases are greater than or equal to $3. Under LePage’s theory, this rebate offer constitutes de facto tying because it induces customers to purchase product Y from firm A when they otherwise might purchase product Y from firm B. In particular, customers may be induced by this offer to purchase product Y from firm A unless firm B reduced the price of product Y by at least $0.20.

3M asserted that firm A in the example has not tied the purchase of product Y to the purchase of product X and that customers remain free to purchase product X and product Y independently, as they did before. Rather, 3M attributed the change in customer behavior to the indirect discount that firm A offers on product Y. Specifically, each customer retains the
ability to purchase product X for $2 (as before) but is now given the additional opportunity to purchase both product X and product Y from firm A for a total cost of $2.80. Thus, the incremental net price of product Y is only $0.80. According to 3M, the $0.20 price reduction required for firm B to retain its customers is the competitive response needed to counter the low incremental price offered by firm A for product Y and not the result of any de facto tie.

**Exclusive Dealing**

Exclusive dealing occurs when a firm requires its customers, or its customers agree, to purchase particular products only from that firm. For example, a pen manufacturer might sell pens to retail distributors only if the distributors agree not to carry pens produced by any other pen manufacturer.

Exclusive dealing arrangements can promote efficiency in numerous ways. For example, exclusive dealing arrangements may ensure that other firms will not share in the benefits of a manufacturer’s promotional efforts without sharing the costs. This type of free-riding could otherwise discourage efficient investment in promotional activities."

Exclusive dealing arrangements are commonplace and generally do not raise antitrust concern. However, exclusive dealing arrangements can be anticompetitive if they persist for a significant period of time and involve a sufficient number of distributors. If a single manufacturer secures long-term exclusive distribution agreements with a sufficient number of distributors, the ability of other manufacturers to compete effectively can be compromised, and competition can be restrained.

LePage’s claimed that 3M restrained competition by securing exclusive distribution agreements with a number of critical distributors. 3M disputed this claim. 3M asserted that there were no instances in which 3M obtained a commitment from any distributor to sell any 3M consumer product on an exclusive basis for any period of time." 3M customers were free to switch suppliers anytime that they chose. They were also free to purchase similar products from multiple suppliers. 3M pointed out that many 3M customers (including all, or virtually all, of LePage’s customers) purchased home and office tape from more than one supplier.

LePage’s countered this argument by alleging that 3M’s rebate programs achieved de facto exclusive dealing by forcing many of 3M’s customers to deal only with 3M in order to obtain rebates on 3M’s alleged monopoly products. LePage’s claimed that the volume thresholds used by 3M to qualify customers for rebate payments were frequently set to ensure

---

"LePage’s expert claimed that exclusive dealing has no justification in terms of consumer or economic welfare.

"3M had agreements with two small distributors that had some elements of exclusivity. 3M claimed that these agreements did not prevent LePage’s from selling private label home and office tape to either customer.

Page 293
that customers would deal exclusively with 3M in order to obtain substantial rebate payments and attempted to show how the private label programs of some customers were influenced by the 3M rebate programs. 3M responded by claiming that, as with the allegation of de facto tying, this was really just a complaint about 3M’s price discounts. If customers chose a single supplier, 3M argued, that was not anticompetitive since those customers might simply be recognizing reduced costs and improved logistics from such an arrangement. In fact, several of 3M’s customers stated that these were precisely the reasons why they limited the number of vendors from which they purchased.

Direct Evidence Regarding Customer Decisions

3M introduced testimony from a number of customers regarding the reasons for their choices among alternative suppliers of home and office tape. Their testimony showed that customers purchased from 3M for many reasons other than the 3M rebate programs, including quality, service, and advertising. 3M argued that this testimony undermined LePage’s claim that customer purchasing decisions were driven by the 3M rebate programs. 3M also pointed to the many customers who continued to purchase from LePage’s despite participating in 3M’s rebate programs, the ability of LePage’s to win new accounts notwithstanding the existence of the 3M rebate programs, and the discounts that 3M continued to provide in response to competition from LePage’s as additional evidence that 3M’s rebate programs did not dictate purchasing decisions or prevent customers from considering LePage’s as a viable alternative supplier of home and office tape.

3M further argued that if its rebate programs had the effect that LePage’s alleged, one would expect to see some correlation between customer conversions to 3M products and the rebates received by those customers. 3M presented evidence showing that there was no such correlation. Specifically, there was no relationship between the year of conversion and the size of the rebate earned in that year for virtually any of the customers who, LePage’s claimed, were foreclosed by the 3M marketing programs. 3M did acknowledge that one or more customers might have altered their purchasing decisions in order to maximize their chances of obtaining a specific growth target and the associated 3M rebate but argued that it was unreasonable to conclude that the rebates systematically played a key role in customer decisions with regard to the purchase of secondary brand and private label home and office tape, given that the rebates were almost never ultimately affected by those decisions.

Efficiency Considerations

An important aspect of evaluating allegedly anticompetitive practices concerns their possible efficiency rationales. LePage’s argued there was no
legitimate efficiency justification for 3M’s alleged exclusionary practices. LePage’s argued that the practices were simply efforts to undermine competition and that if there were any benefits, 3M could have achieved them in ways that did not foreclose competition.

3M responded that its desire to increase the sales and diversification of its products by offering lower prices to its customers was sufficient justification for the conduct at issue. 3M also claimed that it incurred lower costs when dealing with large customers who purchased multiple product lines from 3M. The rebate programs enabled 3M to share these lower costs with its customers in a systematic way. 3M also claimed that the rebate programs helped it to coordinate its marketing strategies across divisions by providing a framework in which the company could speak with one voice to its customers.

3M further noted that the rebate programs responded to the demands of its customers, particularly larger customers who wanted recognition for the volume of business they brought to 3M. Those customers also sought to realize efficiencies from dealing with fewer suppliers, each with broader product lines.

**THE COURT DECISIONS**

The jury in the district court litigation found 3M not guilty of entering into unlawful agreements in restraint of trade and unlawful exclusive dealing but did find 3M guilty of unlawful maintenance and attempted maintenance of monopoly power. The district court judge narrowed the scope of the guilty verdict by ruling as a matter of law that 3M was not guilty of attempted maintenance of monopoly power but upheld the remaining element of the jury decision—unlawful maintenance—and entered a judgment for damages that, when trebled, was over $68 million plus prejudgment interest. 3M appealed the decision to the Third Circuit Court of Appeals.

3M’s argument on appeal was straightforward. Simply put, 3M asserted that the various marketing programs at issue were merely price discounts and that as price discounts the only grounds on which they could be challenged were the predatory pricing standards set forth in the Supreme Court’s *Brooke Group* decision. But, as 3M pointed out, LePage’s did not assert a predatory pricing claim, nor did it assert or attempt to prove that 3M’s prices were less than 3M’s costs, however they were defined. Moreover, 3M said, the jury ruled in favor of 3M regarding LePage’s’ claims of unlawful exclusive dealing and restraint of trade, so there no longer was a legal basis for the guilty verdict with regard to the remaining monopolization claim.

“3M did not attempt to quantify these cost savings at trial.

“That decision held that in order to be predatory, prices must be set below some appropriate measure of the defendant’s costs.
A three-judge appellate court panel heard 3M’s appeal and issued an opinion consistent with 3M’s arguments. The panel affirmed the district court order granting 3M’s motion for judgment as a matter of law with regard to LePage’s attempted maintenance of monopoly power claim but reversed the portion of the district court order pertaining to LePage’s remaining claim of monopolization and remanded the case to the district court with instructions to enter judgment as a matter of law in favor of 3M.

The appellate panel’s decision was subsequently reheard by the full court of appeals en banc. The appeals court ultimately issued an opinion (by a 9-1 majority) reversing the panel’s decision and sustaining the initial jury verdict.

The majority provided a number of bases for its opinion. First, it found that exclusionary conduct (such as exclusive dealing and bundled rebates) can sustain a verdict of unlawful maintenance and attempted maintenance of monopoly power under Section 2 of the Sherman Act. It then rejected 3M’s argument that its marketing programs were nothing more than ordinary price discounts that must be evaluated using the predatory pricing standard established in Brooke Group. Rather, the majority found that the 3M bundled rebate programs could be viewed as de facto tying mechanisms and thus could harm competition by foreclosing competing suppliers from the market.

In addition, the majority distinguished the 3M case from the Brooke Group matter by highlighting the fact that in Brooke Group the allegation involved coordinated activity and not unilateral conduct by a single supplier with monopoly power. Thus, the majority questioned whether the Brooke Group decision was applicable to the facts of the 3M case.

The majority also found that the bundled rebates and other payments made by 3M to its customers could be viewed as methods for implementing exclusive dealing arrangements and thus, like exclusive dealing arrangements, could harm competition by foreclosing competing suppliers from the market. The majority rejected 3M’s argument that because the district court jury did not find 3M guilty of unlawful exclusive dealing agreements or restraint of trade, LePage’s monopoly leveraging, exclusive dealing, and full-line forcing claims were not relevant to the analysis. The majority concluded that the conduct of a monopolist must be viewed in total and not merely by looking at the individual activities in isolation. Finally, the majority highlighted the evidence presented by LePage’s regarding 3M’s intent to “kill” the private label segment and concluded that this evidence was sufficient to overwhelm 3M’s largely undocumented efficiency rationale.

The dissenting opinion offered a very different view of the evidence and the requisite standard of proof. The dissent began by focusing on the testimony of the customers and the many reasons (other than the 3M bundled rebates)...
rebate programs) that were given for customer decisions regarding the purchase of private label home and office tape from 3M rather than from LePage’s. The dissent referred to prior court rulings that volume rebates are legal and pointed out that LePage’s failed to demonstrate that the structure of the 3M rebate programs made it impossible for an equally efficient firm to compete, or even that LePage’s could not have matched 3M’s offer in specific instances. Moreover, the dissent took the view that the failure of LePage’s to make a predatory pricing claim should weaken, not strengthen, its case. The dissent found that even if the rebates implemented exclusive dealing arrangements, they were not anticompetitive because they did not restrict customer choice for an extended period of time. Finally, the dissent concluded that the undisputed fact that the 3M rebates helped 3M to increase its sales provided a valid and sufficient business justification for its conduct, whatever the merits of its other efficiency claims.

FINAL OBSERVATIONS

The dramatically different characterizations of the 3M marketing programs by 3M and LePage’s and the contrasting majority and dissenting opinions of the court of appeals judges regarding the implications of 3M’s conduct highlight the ambiguous state of the law regarding unilateral conduct by firms with dominant market positions. There are strong differences of opinion regarding the extent to which the behavior of dominant firms must be restricted to protect the competitive process and the amount of discretion that should be given to juries in evaluating claims of anticompetitive unilateral conduct.

Some observers, such as those who support the minority court of appeals opinion in LePage’s v. 3M, appear to favor narrowly drawn conduct restrictions and clear standards of proof that competition has been, or is likely to be, harmed by the disputed conduct. Other observers, including those who support the majority court of appeals opinion, appear to take a more expansive view regarding the range of conduct that should be discouraged and the standards that should be applied in establishing harm to competition in specific cases.

In some areas, and particularly with regard to predatory pricing, the courts have made some progress in setting out relatively clear standards for evaluating potentially anticompetitive conduct. Even here, however, there is continuing disagreement, as evidenced by the attempt in the majority court of appeals opinion to limit applicability of the Brooke Group decision to situations in which the defendant does not enjoy significant monopoly power. The differences are even more pronounced when dealing with new or novel forms of conduct. In their joint amicus brief to the Supreme Court recommending that 3M’s request for a hearing should be denied, the U.S. Department of Justice
and the Federal Trade Commission took the view that further academic research and additional judicial experience were needed before the Supreme Court could make a reasonably informed judgment regarding the specific conduct at issue in this case. Since the LePage's decision, a number of economists have examined circumstances under which "bundled pricing" programs might lead to higher (or lower) market prices. Most of these efforts remain fairly specific, so there is still inadequate guidance for antitrust policy toward such pricing schemes. Theoretical work by Kolay et al. (2004), for example, examines "all-units" discounts—that is, a schedule of increasing discounts on all units purchased—and finds they may have either beneficial or adverse effects on consumers. Also from a theoretical perspective, Greenlee et al. (2004) and Greenlee and Reitman (2004) find the possibility of either anticompetitive or pro-competitive effects from bundled discounts but again offer little clear guidance for policy.

In an effort at a more pragmatic approach, Rubinfeld (2005) recounts a number of relevant considerations and then hypothesizes that bundled rebates are anticompetitive if they "(a) reduce consumer welfare, and (b) do so by impairing rivals' ability to make competitive offers." He quickly concedes, however, that this rule does not really offer sufficient guidance to decide particular cases.

As a further indication of the importance of bundled rebates, the recent Antitrust Modernization Commission (AMC), charged with making recommendations with respect to the antitrust laws, devoted a considerable amount of discussion to bundled rebates and specifically to a critique of the appeals court decision: "The decision is too vague and is therefore likely to chill welfare-enhancing bundled discounts or rebates" (AMC 2007, p. 97). The AMC recommended that courts require a plaintiff to demonstrate that the defendant sold the competitive product below its incremental cost, that the defendant would be likely to recoup its losses, and that there would be an adverse effect on competition (AMC 2007, p. 99).

The fact situations that are likely to arise in the future will continue to be rich with nuance and subtlety that cannot be fully captured in simple economic models. "The broader and more fundamental issue concerns the standards that should be employed for establishing that a firm with monopoly power has overstepped the bounds of fair competition and unlawfully excluded rival firms from the market through predatory or other forms of anticompetitive conduct. The Brooke Group decision provides a reasonable starting point with regard to predatory pricing but may need to be modified

"For earlier work, see, e.g., Eaton and Lipsey (1979)."
or clarified if it is to provide a meaningful benchmark for evaluating pricing strategies over the wide range of circumstances that may occur in practice.

It would be similarly useful to develop and articulate standards that can be applied in situations, such as the LePage's v. 3M dispute, in which the claim is that rivals have been excluded from a relevant market by nonprice strategies such as exclusive dealing or tying. Adoption of more general principles, rather than a case-by-case analysis of unilateral firm conduct, would do much to simplify and unify the analysis of claims brought under Section 2 of the Sherman Act.

REFERENCES


CASE 11


John M. Connor

INTRODUCTION

In the evening of June 27, 1995, more than seventy FBI agents simultaneously raided the world headquarters of Archer-Daniels-Midland Company (ADM) in Decatur, Illinois, and interviewed a number of ADM officers in their homes. Serving subpoenas authorized by a federal grand jury sitting in Chicago, the agents collected documents related to ADM’s lysine, citric acid, and corn-sweeteners businesses. Within a day or two, investigators had also raided the offices of four other companies that manufactured or imported lysine. These subpoenaed documents, together with hundreds of secret tape recordings of the conspirators’ meetings and conversations, built a strong case that the five companies had been illegally colluding on lysine prices around the world for three years.

The FBI raids were widely reported in the mass media and unleashed a torrent of legal actions, some of which were still unresolved ten years later. The antitrust actions were the result of an undercover investigation by the U.S. Department of Justice (DOJ) that had begun in November 1992.

By the end of fiscal 1996 (June 30, 1996), ADM was a defendant in an antitrust suit or target of a government investigation in seventy-nine cases, of which twenty-one related to lysine. In subsequent fiscal years, the number of active suits or investigations varied from twenty-two to forty-one (ADM 2001). In 2002, ADM appealed its lysine-conspiracy fine levied by the European Commission, but the European Court of Justice ruled against ADM in 2006.
with the undercover cooperation of ADM’s lysine division president. Within a year of the FBI raids, more than forty civil antitrust suits were filed in federal district courts by direct buyers of lysine. In early 1996, approximately four hundred plaintiffs were certified as a single federal class, and the case called Amino Acid Lysine Antitrust Litigation was assigned to a judge of the U.S. District Court of Northern Illinois. Soon thereafter in April 1996, the three largest defendants offered the federal class $45 million to settle the damages allegedly caused by their price fixing. Final approval of the settlement occurred in July 1996. Additional follow-up suits include about fifteen actions filed by farmers, consumers, and other indirect buyers of lysine in the courts of six states and two Canadian provinces. The lysine cartelists incurred monetary penalties that would total $305 million, of which ADM paid $177 million. ADM was further distracted by shareholders’ suits charging mismanagement by the company’s managers and board of directors.

A few months later, the DOJ sought and obtained guilty pleas for criminal price fixing by the five corporate lysine sellers. Although the corporate members of the cartel pleaded guilty and paid historic fines, not all of the executives who managed the conspiracy were willing to plead guilty. Therefore, the DOJ prosecuted four lysine executives in a highly publicized jury trial held in Chicago in the summer of 1998; three of the four were found guilty and heavily sentenced. The five corporate conspirators were later investigated and fined by the antitrust authorities of Canada, Mexico, Brazil, and the European Union.

The three federal lysine cases were important for at least four reasons. First, they represented the U.S. government’s first completely successful conviction of a global cartel in more than four decades. In the 1940s, the DOJ had obtained convictions of scores of companies that had been

---

1 The two other defendants settled for almost $5 million about a year later.
2 It is difficult to trace most settlements in indirect-purchaser suits, but I estimate that the payout was about $25 to 30 million. The Canadian private damages suit, one of the first such suits in Canada, covered both direct and indirect purchasers and settled for US$3.4 million.
3 The transcript and exhibits of U.S. v. Michael D. Andreas et al. are a major source of primary information on the lysine cartel (Tr.: exhibits from this trial are labeled “Tr. Ex.”). In late 1999, three top ADM officers were sentenced to long prison terms; Andreas got a thirty-six-month sentence, the then-maximum allowed by the Sherman Act (Kanne et al. 1999). One defendant, a managing director of Ajinomoto of Japan, remains a fugitive.
4 The transcript and exhibits of U.S. v. Michael D. Andreas et al. are a major source of primary information on the lysine cartel (Tr.: exhibits from this trial are labeled “Tr. Ex.”). In late 1999, three top ADM officers were sentenced to long prison terms; Andreas got a thirty-six-month sentence, the then-maximum allowed by the Sherman Act (Kanne et al. 1999). One defendant, a managing director of Ajinomoto of Japan, remains a fugitive.
5 The two other defendants settled for almost $5 million about a year later.
6 It is difficult to trace most settlements in indirect-purchaser suits, but I estimate that the payout was about $25 to 30 million. The Canadian private damages suit, one of the first such suits in Canada, covered both direct and indirect purchasers and settled for US$3.4 million.
7 The transcript and exhibits of U.S. v. Michael D. Andreas et al. are a major source of primary information on the lysine cartel (Tr.: exhibits from this trial are labeled “Tr. Ex.”). In late 1999, three top ADM officers were sentenced to long prison terms; Andreas got a thirty-six-month sentence, the then-maximum allowed by the Sherman Act (Kanne et al. 1999). One defendant, a managing director of Ajinomoto of Japan, remains a fugitive.

---

The five corporate conspirators were later investigated and fined by the antitrust authorities of Canada, Mexico, Brazil, and the European Union. The three federal lysine cases were important for at least four reasons. First, they represented the U.S. government’s first completely successful conviction of a global cartel in more than four decades. In the 1940s, the DOJ had obtained convictions of scores of companies that had been
members of international cartels that had operated between the two world wars (Wells 2002). These suits had been initiated in the late 1940s and had been completed by around 1950. During 1950-1995, the DOJ had attempted to prosecute only three international cartels. In all three cases, the DOJ failed to prevail at trial either because essential evidence located abroad could not be subpoenaed or because juries would not convict U.S. business executives on the testimony of their foreign coconspirators. The criminal lysine convictions signaled to would-be price fixers that the DOJ could win convictions of international schemes at trial.

Second, the conviction of the lysine cartel was the first public manifestation of a sea change in enforcement priorities by U.S. and overseas antitrust officials. Prior to 1995, less than 1 percent of the price-fixing indictments by the DOJ involved at least one non-U.-S.-based corporation or non-U.S. resident. By contrast, since 1998 more than half of all criminal price-fixing indictments have been brought against non-U.S. conspirators (Connor 2007a, fig. 1.1). The investigation of the lysine cartel led directly to the discovery and successful prosecution of thirty multinational corporations that participated in global price fixing in the markets for lysine, citric acid, sodium gluconate, and ten bulk vitamins. Since 1996, nearly two hundred international cartels have been uncovered and prosecuted by the DOJ, the Competition Policy Directorate of the European Commission (DG-COMP), and other antitrust authorities around the world (Connor and Helmers 2006). Cartel enforcement remains a high priority for the Antitrust Division of the DOJ, which is devoting about 30 percent of its resources to criminal price-fixing prosecution.

Third, the lysine cartel case demonstrated the government's intention to employ tough, "blue-collar" criminal investigative techniques to what had been formerly been treated as gentle, "white-collar" misdemeanors. In the three-year investigation that preceded the FBI's 1995 raid, the DOJ showed that it was prepared to use all of the tools of its profession that it habitually employs in gathering evidence against drug cartels and other forms of organized crime, including videotaping and obtaining the cooperation of foreign

"Prosecutions of the uranium cartel (1978) and industrial diamonds (1994) were hampered by the absence of witnesses and documents that were outside U.S. jurisdiction. The DOJ lost at trial when prosecuting one member of the thermal fax paper cartel in 1995. Finally, although the DOJ was victorious in a U.S.-Canada cartel that fixed the prices of plastic dinnerware (1996), that cartel was only tangentially international in scope. In contrast, the three ADM executives were sentenced to a collective total of ninety-nine months in prison.

"Joshua (2006, p. 16) argues that the European Commission had had its epiphany about the ubiquity and harmfulness of hard-core international cartels in the late 1980s.

"In fiscal 2002, the division planned on allocating 258 full-time equivalent positions (FTEs) to this activity (DOJ 2002). However, if one adds the resources of U.S. attorney offices and the FBI, about three thousand FTEs were devoted by the U.S. government to fighting cartels, at an annual cost of $440 million.

"Copies of the lysine videotape and transcripts are available at no charge by mailing a request to: U.S. Department of Justice, Antitrust Division, Freedom of Information Act Unit, 325 Seventh Street, NW, Suite 200, Washington, D.C., 20530.
police organizations. During guilty-plea negotiations with targeted cartel conspirators, prosecutors have made deft use of a wide range of possible penalties to instill cooperation, including threatening crippling fines, imposing significant prison sentences, and perpetually barring convicted felons from entering U.S. territory.

Finally, the lysine cases and those that followed soon thereafter showed that although the penalties for criminal price fixing had escalated enormously, they may still be inadequate to suppress recidivism.

Congress raised the maximum corporate statutory fine for Sherman Act violations to $10 million in 1990 and to $100 million in 2004. U.S. Sentencing Guidelines first promulgated in 1987 permit government fines of up to 80 percent of a guilty company's affected sales or, alternatively, criminal penalties as high as "double the harm" caused by a cartel (Connor and Lande 2005). That is, corporations can be fined twice the monopoly overcharge generated by a cartel, an amount that can exceed the statutory cap when percentage overcharges are high (Connor 2007a, pp.77-80). ADM, the leader of the lysine cartel, was fined $100 million for its role in two criminal price-fixing schemes—a record amount that was twice eclipsed in the late 1990s by leaders of highly injurious global cartels. Since the lysine cartel was fined, worldwide antitrust penalties on international cartels have exceeded $25 billion (Connor and Helmers 2006).

U.S. government fines are mere paper cuts compared with the financial wounds that may be inflicted by plaintiffs in civil actions. Direct buyers suing in federal courts, the principal focus of this chapter, are entitled to seek treble damages. However, antitrust liability does not stop there. More than twenty-five states allow their residents who are indirect purchasers to sue in state or federal courts, most of which permit treble damages. State attorneys general increasingly have banded together to pursue antitrust claims in federal courts (parses patriae suits) to recover treble damages for their state governments and for corporate and individual indirect buyers residing in their states. Not counting the losses associated with derivative shareholders' suits, legal fees, and reputational effects, corporations accused of criminal price fixing now face maximum U.S. antitrust liabilities that could range up to seven times the cartel's U.S. overcharges. The fines and prison terms meted out to cartel managers have risen also.

"ADM was the second firm to be fined above the $10 million statutory limit but the first to be widely reported by the business and popular media. In 1999, Hoffmann-La Roche paid $500 million for its participation in six global cartels for bulk vitamins. Further details on global cartel sanctions are given in Connor (2004).

"In October 2000, the attorneys general of more than forty states announced a settlement totaling $340 million to be paid by the six largest members of the vitamins cartels (Connor 2007a, p. 499).

"This is a theoretical upper limit. In practice, cartels rarely incur penalties that exceed treble damages (Connor 2006).

"The median individual fines are below $100,000, but a few have been high. The employer of one German CEO, the ringleader of the graphite-electrode cartel, paid a U.S. fine of $10 million to
The major role played by economic analysis in horizontal price-fixing cases is the calculation of the overcharge on buyers in markets affected by a cartel. The overcharge is the value of purchases of a cartelized product actually made, minus what the sales would have been for the same volume of product absent the cartel. Accurate estimates of conspiracy-induced overcharges are primarily of importance to determine the recovery of civil damages but also can be the basis for the calculation of U.S. government fines. In criminal price-fixing cases, the U.S. Sentencing Guidelines obligate federal prosecutors to seek corporate fines that are based on either a defendant's affected sales or its overcharges. Fines up to $25 million can be assessed on individual price fixers from a sliding scale that depends on the size of the overcharge. In summary, both corporate and personal penalties for price fixing are, in principle, closely related to cartel overcharges.

U.S. price-fixing penalties (and those of most other nations) are based on the legal-economic theory of "optimal deterrence." A simple version of that theory states that fines and private damages paid should be equal to the economic harm caused divided by the probability of detection and prosecution. Such penalties will reduce the rate of formation of cartels—if not to zero, then to an ideal trickle. The probability of catching secret cartels is widely believed to be between 10 percent and 33 percent avoid a prison sentence of about six months. In 2002, the chairman of auction house Sotheby's was fined $7.5 million for price fixing.

For a broad historical analysis of cartel overcharges, see Connor and Lande (2005). Prior to 1974, the maximum prison sentence was one year. When price fixing was made a felony in 1974, the maximum sentence rose to three years, and in 2004 an amendment to the Sherman Act raised the cap to ten years (AMC 2007, p. 296).

By long-standing custom, the DOJ brings 90 percent or more of all horizontal price-fixing indictments as criminal matters, and only hard-core cartels are prosecuted criminally (AMC 2007, p. 297).

"The guidelines were mandatory from 1987 to January 2005, at which time following the Supreme Court's decision in U.S. v. Booker, 543 U.S. 220 (2005), they became voluntary (Connor and Lande 2005). The guidelines are complex in their details (see http://www.ussc.gov/guidelin.htm) but are intended to deter price fixing by imposing punitive penalties.

There are two approaches. If prosecutors believe that the overcharge is around 10 percent of affected sales, a base fine of 20 percent of sales is calculated, and the base fine is raised by a complex list of factors that yields a numerical culpability multiplier between 0.15 and 0.40. This method is intended to be a convenient proxy for an overcharge-based fine calculation. Alternatively, if prosecutors have evidence that the overcharge was significantly above 10 percent or want to impose a fine above the Sherman Act limit, then the maximum fine for criminal price fixing is double the cartel's overcharge. In practice, one of the two methods produces a proposed fine that becomes the starting point for DOJ-defendant plea bargaining; guilty-plea negotiations typically result in actual fines well below the starting point.

In rare published versions of the details of an antitrust plea-bargaining episode, DOJ officials lowered the fine demanded of ADM in stages from as much as $400 million (Lieber 2000, p. 36) to $100 million (Eichenwald 2000, pp. 487, 507-511, 521-523). The number of ADM officers to be indicted also was reduced.

Harm is customarily equated with the overcharge but should also include the deadweight loss. Ideal in the sense that the expected private costs (e.g., penalties, antitrust training programs, etc.) and costs of the justice system will be equal to the cartel's expected profit benefits.
Thus, the optimal penalties for cartels are equal to 300 percent to 1000 percent of the cartel’s overcharge.

The primary purpose of this case study is to illustrate the computation of overcharges in a forensic setting with facts drawn from the lysine cartel of 1992-1995. Most of the issues regarding such calculations arose during preparations for the private federal class action, *In re Amino Acid Lysine Antitrust Litigation*, in the summer of 1996. However, as additional time and economic data became available, more refined estimates of the lysine-cartel overcharge became available for the sentencing phase of the criminal case, *U.S. v. Michael D. Andreas et al.* This study also illustrates the difficulties of imposing corporate monetary penalties in international criminal price-fixing episodes that are capable of optimally deterring cartel formation.

**THE ECONOMICS AND LAW OF CARTELS**

A hard-core cartel is an association of two or more legally independent businesses that explicitly agree to coordinate their prices or output for the purpose of increasing their collective profits. Some cartels are organized by state agencies or government-owned corporations; other cartels have been formed by multilateral treaties to attempt to smooth commodity price cycles. This chapter is concerned only with private business cartels that operate unprotected by the cloak of national sovereignty.

Economics views cartels as a special type of oligopoly, an extralegal joint venture of businesses that are normally rivals in the same industry. The goal of a cartel is usually to increase the joint profits of its members to a level as close as possible to that of a monopolist; the strategy of a cartel is to implement one or more of the “restrictive business practices” popularly known as price fixing. A key feature of private cartels is that the participants must forge a “contract” that is self-enforced. Cartels almost always explicitly contract to raise their list prices, to lower total production, or both; they may also reinforce this basic agreement by fixing market shares for each member, allocating specific customers, imposing uniform selling conditions, sharing sales information, monitoring price agreements, pooling and redistributing profits, adopting a method for punishing deviants, and hiding or destroying evidence. The time and management resources required to negotiate the formation of a cartel and to carry out the agreements can be substantial.

Economic models of cartels emphasize the necessity of high concentration and of product homogeneity in an industry (Stigler 1964, Dick 1998, Connor 2007a). Without small numbers of member-sellers and reasonably standardized products, the transactions costs of forming and maintaining a group consensus would become too high relative to the anticipated increase in profits. Moreover, because cartel agreements must be self-enforced and
because there is always a profit incentive for cartel members to cheat on the cartel’s agreement (i.e., to sell more or at a lower price than that agreed upon), only small numbers and homogeneity will keep the information costs of detecting cheating within acceptable bounds. Other conditions that facilitate the formation and stability of cartels include a large number of buyers, a small amount of noncartel production capacity, equality of production costs across firms, and relatively stable or predictable demand conditions. High barriers to entry into the industry will facilitate the formation and longevity of cartel agreements.

Section 1 of the 1890 Sherman Act deems cartels illegal, and since 1927 (United States v. Trenton Potteries Company et al., 273 U.S. 392) they have been per se illegal. That is, an explicit agreement to fix prices is an illegal “conspiracy in restraint of trade,” regardless of the agreement’s actual impacts on market prices or output. Outside the United States, the competition laws of most industrialized nations judge the illegality of a cartel under a rule of reason. In practice, however, non-U.S. competition-law agencies routinely prosecute all of the naked cartels that they discover. In the EU rare exceptions are made for cartels with significant benefits for consumers from technological innovation. Many countries, the United States included, permit registered cartels to coordinate exports.

Strict enforcement of laws against overt price fixing is a public policy widely supported by economists and legal scholars of all stripes. They may differ as to the causes giving rise to collusive behavior and as to the likelihood of long-term success, but they are unified in their evaluation of the negative economic effects of cartels. Effective cartels cause unrecoverable losses in production and consumption, transfer income from customers to the stakeholders of cartel members, and often engage in wasteful rent-seeking expenditures (Posner 2001, ch. 1).

INDUSTRY BACKGROUND

Lysine is an essential amino acid, a building block for proteins that speed the development of muscle tissue in humans and animals. Food derived from animal and marine sources normally provides humans with sufficient lysine to ensure healthy muscle development. Certain vegetables, soybeans in particular, are also good sources of amino acids; expensive pharmaceutical-grade or “natural” lysine is chemically extracted from vegetable matter.

"Enforcement in Canada mimics the per se rule, and since the 1980s the EU has evolved toward a U.S.-style, conspiracy-based theory of cartel enforcement (Joshua and Jordan 2004). Other cartel-like behavior, such as patent pooling or standards setting, may be prosecuted in the U.S. as civil violations of the Sherman Act and be judged under the rule of reason.
In 1956, scientists in Japan discovered that amino acids can be produced as a byproduct of bacterial fermentation (Connor 1999). By 1960, two Japanese companies, Ajinomoto and Kyowa Hakko, were selling commercial quantities of lysine utilizing these new biotechnologies. From the beginning, lysine was produced by fermentation at a far lower cost than chemically extracted lysine; continuing improvements in production technologies have brought the cost of so-called synthetic lysine down to less than one-fifth that of pharmaceutical lysine. The lower prices of lysine in turn made it cost effective to incorporate manufactured lysine into animal feeds. Today, well over 90 percent of the world’s lysine supply is made by biotechnology and is used as a supplement in animal feeds, principally for swine, poultry, and aquaculture.

Evolution of the Industry

At prices generally around $1 to $2 per pound, worldwide demand for lysine by animal-feed manufacturers soared from nothing in 1960 to almost 70 million pounds in 1980. In the 1980s, global consumption of lysine grew by 16 percent per year; in the 1990s, volume growth was still a heady 12 percent annually (Connor 2007a, ch. 7). In the early 1990s, approximately two-thirds of the demand for lysine originated in North America and western Europe, areas with the highest levels of meat consumption and with consumers most willing to pay for lean meats.

The Japanese duopoly initially satisfied global demand by exporting from its two domestic plants. Ajinomoto made the first move abroad by building a large plant in France in 1974. Kyowa Hakko opened its first overseas lysine plant in Mexico in 1980 and its second in Missouri in 1984. Ajinomoto, which had about twice the capacity of Kyowa Hakko, responded by opening its own U.S. lysine plant in Iowa in 1986. After they were in operation, the Japanese firms implemented significant capital expansions to their plants outside Japan every two or three years.

Lysine was a Japanese duopoly until 1980, when the South Korean conglomerate Sewon opened a new plant in its home country. Sewon never expanded through direct investment abroad, instead relentlessly expanding its sole plant and exporting most of its output to Asia and Europe. Sewon reached its goal of achieving a 20 percent world market share by the late 1990s, but at the cost of massive borrowing.

The lysine oligopoly colluded successfully at least three times prior to 1992 (Connor 2007a, pp. 174-175). It fixed prices in Japan in the 1970s and 1980s and in Europe in the 1980s. From 1986 to 1990, Ajinomoto and Kyowa Hakko fixed prices and divided the U.S. lysine market 55:45 percent. The U.S. price of lysine at times reached just over $3 per pound in the late 1980s. In brief, when the lysine biotech industry consisted of two or three Asian producers, collusive behavior was more often the norm than was uncooperative or classic competitive behavior.
Entry in the 1990s

Patents on high-yielding microbes and technological secrecy largely prevented new firms from enjoying the high growth and large profits being made in the lysine industry. Leading French and German biotechnology firms attempted to form lysine joint ventures in the 1970s and 1980s but were thwarted from doing so (Connor 2007a, pp. 175-176). With two exceptions, only very small-scale entry occurred in the 1990s (Connor 2007a, pp. 180-182).

In early 1991, two newcomers turned the lysine industry into a five-firm oligopoly. ADM, according to a plan finalized in late 1989, opened the world’s largest lysine-manufacturing plant at its headquarters in Decatur, Illinois, in February 1991. Within eighteen months, ADM’s plant had expanded global production capacity by 25 percent above year-end 1990 levels, and by 1993 ADM’s single plant accounted for one-third of global capacity (780 million pounds). ADM’s strategic objective was to acquire a global market share equal to the industry leader, Ajinomoto. Ruthless price cutting by ADM and the sudden appearance of large excess capacity caused lysine prices to plunge 45 percent in the first eighteen months of the Decatur plant’s operation. Global oversupply was exacerbated by the simultaneous opening of a smaller plant in Indonesia controlled by the South Korean food firm Cheil Sugar Co. ADM’s aggressive entry into the lysine industry was the precipitating event in the formation of a new lysine cartel in 1992.

Cartel Behavior 1992-1995

Ajinomoto, Kyowa Hakko, and Sewon began meeting as early as April 1990 to try to forge a plan to cope with ADM’s entry, but they were fatalistic about ADM’s impending success. After ADM entered production, the Asian manufacturers repeatedly signaled their willingness to raise lysine prices, but ADM appeared to be steadfast in its drive toward sharing global dominance. By mid-1992, ADM had captured an impressive 80 percent of U.S. sales, and it was exporting more than half of its production. Ajinomoto and Kyowa experienced large operating losses in late 1991 and early 1992. In June 1992, the U.S. transaction price bottomed out at $.68 per pound, which was $1.10 per pound below the long-run marginal cost of ADM (see the discussion of costs below).

By early 1992, the Asian incumbents were considering asking ADM to join them in a more cooperative arrangement. It must have seemed something of a godsend when in April 1992 the president of ADM’s lysine division and another more senior ADM officer showed up in Tokyo to

---

Details may be found in Connor (2000, app. A; 2001a, ch. 8). Many of these facts were corroborated by testimony or exhibits from the 1998 criminal trial. For popular accounts of the cartel, see Eichenwald (2000) and Lieber (2000). The final U.S. legal decision was the opinion of the U.S. Court of Appeals for the Seventh Circuit in U.S. v. Michael D. Andreas and Terrance Wilson, 216 F. 3d 645 (2000).
propose the formation of a lysine "trade association." Under the guise of establishing such a trade group, Ajinomoto, Kyowa, and ADM officers met in Mexico City in June 1992. This was the first of twenty-five multiparty price-fixing meetings among the five corporations that joined the cartel; dozens of supplementary bilateral meetings by regional sales managers and hundreds of telephone calls cemented agreements on prices in as many as thirteen currencies. The price agreements covered only dry feed-grade lysine. In early 1993, a brief price war broke out among the conspirators, mainly because of ADM’s insistence that the participants had to agree to global market shares. After a top-level meeting in October 1993 resolved the issue, the cartel displayed a high level of harmony and consensus. Cheating was restrained in part by largely accurate monthly reporting of each company’s lysine sales volume to all the members of the cartel.

The lysine cartel ended with the FBI raid on cartel offices in June 1995, almost exactly three years after the first price-fixing meeting had occurred. During that time, the average U.S. transaction price of lysine (manufacturers’ delivered price) rose from $0.68 per pound when the cartel began operating to a plateau of $0.98 (October-December 1992), fell again to $0.65 (May 1993), and rose quickly again to above $1 for most of the remainder of the conspiracy period (see Figure 11-1). Prices in the European

"A year or two later, the International Amino Acids Manufacturers’ Association was formed and recognized as a “working party” of the Agriculture Directorate of the European Commission.

"In the U.S. market ADM sold a somewhat diluted aqueous version delivered in tanker trucks to nearby customers. On an active-ingredient basis, liquid lysine was less expensive but highly correlated in price movements to the powder form. Liquid lysine accounted for well under 5 percent of the U.S. market.

Page 309
Union closely tracked those in the United States, albeit at a level $.10 to $.25 higher. "Target prices were also higher than the U.S. target price in Latin America, Japan, Oceania, and most parts of Asia. However, for the rest of this chapter, only U.S. prices will be analyzed.

The Costs of Collusion

As mentioned above, there is considerable sentiment among some econo­mists that the costs of forming and maintaining a collusive contract are so high that the incidence of cartels is low and their lives fleeting." The his­tory of the lysine cartel and related global cartels prosecuted in the late 1990s does not support this sanguine view.

Internal memoranda and extensive trial testimony by cartel partici­pants confirm that the conspirators reasonably anticipated that the rewards from price fixing would far outweigh the costs of operating the cartel (Connor 2007a, ch. 8). At a key meeting in late 1992, a top ADM official expressed the expectation that their recently concluded agreement would generate $200 million in joint profits annually in a global market for lysine that generated from $500 to $700 million in annual sales. His prediction, from ADM’s perspective, was accurate: ADM did earn approximately $200 million in profits from the cartel over three years with its one-third share of sales in the worldwide lysine market. Direct operating costs of the cartel were modest."

"The correlation in prices is even higher when one compares the U.S. price in dollars with the EU price in Deutschmarks. That is, that USS/DM exchange rate, which is rather unpredictable, introduced more variability into the European price because the conspirators used the dollar to fix prices quarterly.

An accessible treatment of the inherent instability of cartel agreements may be found in Posner (2001, pp. 60-69). Among the obstacles to agreement are an own-price elasticity of demand that is too high at the competitive price, uneven costs of production among potential cartel members, product heterogeneity, a steeply rising marginal cost curve, a large fringe of suppliers that will not join the cartel, lower costs of production by the fringe producers, and the difficulty of apportioning reductions in output among the cartelists. After the cartel is formed each member has an incentive to cheat by either cutting price or offering an improved product; cheating is difficult to detect; and virtually the only way to punish deviants is through expensive, self-destructive trigger mechanisms such as price wars.

"During the four years of preliminary negotiations and actual cartel operation, each of the four (later five) companies sent two men to meetings held on average once every three months. Late in the conspiracy, regional sales managers became involved, but the total number of conspirators never exceeded forty (Connor 2000, app. A). Counting the monthly production reports submitted by each firm and other communications, it appears that each corporate member of the cartel man­aged the conspiracy with an input of fifteen to twenty-five person-days per year. Total labor costs for all corporate conspirators could not have exceeded $1 million for the entire conspiracy period.

Neither is there any reference to cartel coordination of cost-saving production technologies or capital investments. What is known is that ADM’s president was involved in illegal industrial espi­onage that stole lysine-production secrets from employees of Ajinomoto’s plant in Iowa (Connor 2007a, pp. 212-213). Indeed, the technological leader in fermentation techniques, Ajinomoto, later sued ADM for patent infringement concerning a genetically modified microbe that yielded more lysine per unit of dextrose. Moreover, the pell-mell expansions of Sewon’s and Chief’s..."
It is certainly true that the cartel members squabbled frequently and that the two smallest members, both South Korean companies, were strongly inclined to cheat on the price and market-share agreements. Infighting led to one sharp price war for a few months in early 1993. However, a number of techniques adopted by the cartel and the impressive diplomatic skills of the cartel’s dual leaders, ADM and Ajinomoto, kept the effects of cheating to tolerable levels.

Among the most important practices that cemented cartel harmony was the tonnage quotas agreed upon in late 1993. Combined with accurate monthly sales reports and politic concessions of additional quotas to the two Korean firms, the market-share agreements were honored with impressive precision throughout 1994 and 1995. The formation of an amino acid trade association under European Commission sponsorship provided excellent cover for the group’s illegal meetings in Europe and elsewhere. A compensation system was adopted to punish members that exceeded their quotas, but it was never necessary to implement the system. ADM, with its new efficient plant and ample excess capacity, frequently reminded the cartel of its willingness to flood the market with lysine; its threats were credible because it had twice driven the world price of lysine to below its own average total cost of production, inflicting the others with operating losses. Moreover, ADM had taken the rare step of inviting its rivals in the lysine market to an intimate tour of its capacious production facilities. Finally, it should be recalled that the three largest Asian companies in the 1992-1995 cartel had had a great deal of experience in organizing price-fixing schemes for two decades. ADM, too, is a serial price fixer.

MEASURING THE OVERCHARGE

The monopoly overcharge is the difference between what a buyer has paid for a cartelized product and what a buyer would have paid absent the cartel. Under the U.S. antitrust laws, a successful plaintiff is entitled to treble the dollar overcharge, which is then multiplied by the number of units purchased.” Two features of estimating cartel damages help simplify the analytical task. First, the market-definition problem, which is so critical in monopolization and merger cases, is usually not an issue. For effective hard-core conspiracies, cartelists self-define the appropriate product and

plants were a constant source of friction at cartel meetings (Connor 2007a, p. 214). Nearly all empirical investigations of modern private cartels fail to find evidence of static cost efficiencies due to cartel coordination (Audretsch 1989).

“Besides lysine, since 1990 ADM was penalized for conspiring to fix prices in eight other markets: carbon dioxide gas, sodium gluconate, citric acid, high fructose corn sweeteners, corn dextrose, corn glucose, monosodium gluconate, and other nucleotides (see Connor and Helmers 2006).

“Equivalently, one may compute the percentage increase in price for each time period during the conspiracy and then multiply these percentages by the value purchased in each period.
geographic market.” Second, the time period for intended cartel price effects is usually an admitted fact. Actual market price changes often will lag by several weeks the starting date and by several months the ending date. It is possible that the cartels achieved no pricing power in the market. However, in most cases the collusive period is treated as parametric information.”

Information on actual transaction prices and quantities sold is usually readily available from the parties in such cases, but the unobserved “but-for” price must be inferred using economic reasoning.” Enter the economists.

Methods of Calculation

There are five generally recognized methods of calculating an overcharge (Page 1996, Hovenkamp 1998). Proving an antitrust injury in U.S. courts depends on the preponderance of the evidence in the case, but the amount of damages is decided according to a lower standard: that of reasonableness. Reasonableness usually entails a “formula” (a precise method of calculation) that can be applied to data likely to be available to the analyst. Each of the five methods of computation below will meet the legal standard of reasonableness, and forensic economists often will examine more than one method to see if they are mutually supporting.

Three methods were applied to calculate the lysine cartel’s overcharge in the U.S. market. One of them, the before-and-after method, was employed by economists acting as experts for the two sides in the civil treble-damages suit. Opinions and rebuttals were exchanged during May–July 1996 prior to a fairness hearing for the federal class of plaintiffs and after the defendants provided their average monthly prices for 1991–1995. Two other methods of analysis, one using costs of production and the other a time-series

“Under U.S. law prior criminal convictions are prima facie evidence for follow-on civil damages cases. Market definition could become a forensic issue if the alleged agreement must be proven from circumstantial rather than direct evidence. Market definition will also require analysis if the alleged cartel conduct was a facilitating practice rather than an overt agreement on price or output.

“In the lysine case, the length of the collusive time period was an issue that divided experts for the two sides (Connor 2001). It is usually uncontroversial to use the conspiracy period for the affected sales period, but in the lysine case the defendants had not yet agreed to plead guilty. Consequently, the opt-out plaintiffs had to depend on press reports to define the affected period. Plaintiffs chose August 1992–December 1995 to be the affected sales period. In fact, court testimony later revealed that the lysine cartel had two episodes divided by a brief price war in the spring of 1993. Defendants’ experts decided that only the second episode was effectively cartelized. White (2001) refers to an “unusual and suspicious” pattern of “uncharacteristic stability” in lysine transaction prices from September or October 1993 to February or March 1995 (see Figure 11-1).

“Single damages under the law in most court circuits are precisely equivalent to the income transfer due to the exercise of market power. Single damages are slightly higher than the stream of monopoly profits accruing to the cartel members because operating the cartel requires the expenditure of some management resources. In some circuits, the deadweight loss may be permitted as an additional source of damages (Page 1996). Some legal theorists argue that a buyer’s lost profits are a conceptually superior measure of damages (Hovenkamp 1998, p. 658).
econometric market model, were carried out in 1999 with the benefit of data from exhibits filed in the Chicago criminal trial.

The Before-and-After Method

This method has been used to calculate antitrust damages in U.S. civil cases since at least the 1920s (Hovenkamp 1998, p. 661). "Before-and-after" is something of a misnomer because the "before" period is really any non-conspiracy period—whether before, after, or during an intermediate pause in price fixing. It is important that the "before" period be one that is quite comparable with the conspiracy period with respect to demand and supply conditions. Shifts in buyer preferences, the appearance or disappearance of substitutes, or changes in the cost of production of the cartelized product during the affected period can cause overstatement or understatement of the overcharge.

Choosing the months to employ for the "before" price requires judgment. If real prices were fairly constant for one to three years before a cartel began, then averages of these one, two, or three years' prices might serve well as benchmarks. One problem often encountered is that cartel formation frequently occurs after a recessionary period in the industry; if so, then the benchmark price might be understated and the damages overstated. Similarly, a predatory episode prior to cartel formation could result in an overestimate of the overcharge, as happened in lysine. On the other hand, older precartel prices could have been generated by a previous cartel episode unknown to plaintiffs. This seems to have been the case in the vitamins cartels in 1985-1989 (Bernheim 2002, fig. 8-2).

Similarly, postcartel prices can yield equally problematic benchmarks. Ideally, conduct may be a reversion to noncooperative equilibrium. However, defendants may pursue a strategy of unsustainably low pricing to mollify their angry customers. Alternatively, the cartel experience may allow its participants to form more stable conjectures than were possible before the cartel existed, in which case even uncooperative conduct inflates postcartel prices as a carry-over effect of the earlier cartel behavior. Harrington (2004) developed a model in which former cartelists keep prices high in order to reduce their liabilities in follow-on civil suits.

This method lends itself to simple graphical treatments. Most commonly, the analyst fits either a straight line from the "before" price forward in time until the end of the cartel or a line backward from the "after" price to the beginning of collusion. A slightly more sophisticated approach is to draw a straight line from the before price to the after price. If it has an upward slope, costs probably increased and vice versa.”

“If this line has a slope, this may be a rough way of accounting for changing costs, technological progress, or a secular shift in demand during the affected period. Bernheim (2002, fig. 12-1) illustrates this approach in a damages analysis prepared for plaintiffs in the U.S. vitamin E cartel suit.
In April 1996 ADM, Ajinomoto, and Kyowa offered the federal class of lysine direct purchasers (about four hundred companies) $45 million to settle the suit. This offer came at a time when the DOJ’s criminal investigation appeared to be stalled. Indeed, a rather unusual feature of the civil suit is that the settlement offer was made four months before the government obtained the first of its guilty pleas. In the majority of criminal cartel cases, treble-damages suits are follow-on actions that are settled out of court or go to trial well after guilty pleas are made in a government case—pleas that are by law prima facie evidence in the following civil actions (Lande and Davis 2006). Moreover, civil plaintiffs can benefit from facts admitted in the pleas (e.g., conspiracy dates). In this case, through discovery plaintiffs were provided with only two slim bits of information: average U.S. monthly selling prices of lysine for the years 1990-1995 and annual sales of the four largest sellers (Cheil Sugar did not participate in discovery). From public sources the only other potentially useful data were list prices of lysine, international trade in lysine (value and volume), and U.S. prices for corn and soybeans. Changes in corn prices drove most of the variability in the cost of manufacturing dextrose, the principal feedstock and largest input for making lysine."

The decision faced by the judge in July 1996 was whether the $45 million proposed settlement was fair and reasonable (Connor 2007a, pp. 394-399). The proposed settlement had been negotiated over a period of about three months by the lead class counsel and the law firms representing the lysine makers. Many of the larger lysine plaintiffs were dismayed at the small size of the award and what they perceived to be unassertive legal representation. They had to weigh two options: (1) stay in the class and take a riskless three cents on each dollar of lysine purchased or (2) opt out of the class and face the uncertainty of either a higher settlement or nothing at all. Naturally, it was in the interests of the potential opt-outs to persuade the judge to reject the proposed settlement.

The main issues with respect to calculating the overcharge were the length of the time period of the affected period and the "but-for" price. It is conventional to use the conspiracy period for the affected sales period, but recall that the defendants had not yet agreed to plead guilty. Consequently, the opt-out plaintiffs had to depend on media reports that the conspirators had first met in June 1992 and had continued colluding until the FBI raid

"Dextrose was the foundation for fermentation in Ajinomoto’s and ADM’s U.S. plants; Kyowa Hakko’s Missouri plant used sucrose. Dextrose accounted for 38 percent of variable costs of manufacturing lysine in ADM’s plant and 32 percent of total manufacturing costs (Connor 2007a, p. 229).

"In the interests of expediency, the judge had awarded the role of lead class counsel on the basis of a fixed-fee auction that provided the class counsel with "... little incentive to maximize the recovery for the class" (Coffee 1998, p. B6). The fee was capped at $3.5 million for any settlements above $25 million. The firm hired no economists to analyze the overcharge issue. The legal fees, at 7 percent of the settlement, were very low by historical standards; normally fees are 20-30 percent of recovery in class actions for price fixing."
in June 1995. Because there appeared to be lags between the time that the cartel set a list price and the time that the transaction price fully responded, August 1992-December 1995 was chosen to be the affected sales period.\textsuperscript{37}

The hypothetical nonconspiracy benchmark price was the most contentious judgment that had to be made. Given the paucity of price observations (only seventy-one months), three periods seemed to be the leading candidates: (1) average prices prior to August 1992; (2) a nadir in prices in mid-1993 caused by a disciplinary price war; and (3) prices after June 1995. Ideally, but-for prices should be long-run equilibrium prices, averaged over fairly long periods of about one to three years. However, prices from the February 1991-July 1992 period were affected by an earlier lysine cartel as well as ADM’s massive entry. Not only was market structure shifting, but also costs were changing because of ADM’s learning-by-doing. ADM’s new plant suffered a number of contamination incidents during its first year but very few thereafter. That is, most of the early precartel period appeared to be in disequilibrium.

Like the summer of 1992, the summer of 1993 seemed to be a return to a regime that exhibited highly competitive pricing conduct; journalistic sources, later confirmed by memoranda of the cartel's meetings around this time, reveal that bickering among cartel members resulted in a return to aggressive pricing behavior (Connor 2007a, pp. 213-216). When they later became available, ADM’s production records showed a surge in output in early 1993 that probably triggered the crash in lysine prices. In consideration of these factors, the opt-out plaintiffs chose May-June 1992 and April-July 1993 as the but-for periods. Perhaps accidentally, prices averaged $.70 per pound in both periods.

However, the third candidate period did not appear to be useful for competitive benchmarking. First, only six data points (months) were available for the postcartel period, and the effects of cartel behavior might well lag for several months after the cartel was exposed. Second, the shadow price of lysine had forced the cartel to drop its prices in early 1995, but then, just after the FBI raid, prices climbed precipitously for the rest of 1995.\textsuperscript{38} One might hazard that the former cartel members had learned how to collude tacitly by following movements in the shadow price, because other demand and supply factors did not seem to explain the late-1995 rise in prices.

\textsuperscript{37}Lags were created by thirty-to-forty-five-day price protection clauses in most sales contracts, by delivery chains (particularly overseas deliveries), and information lags. In fact, the observed lags were mostly between two to four months except for minimal spot sales. The lags appeared to be asymmetric: longer when prices were declining and shorter when responding to upward changes in list prices by the sellers.

\textsuperscript{38}The shadow price of lysine was governed by a technical rule-of-thumb followed by the animal-feed industry: Three pounds of lysine and ninety-seven pounds of corn were nutritionally equivalent to one hundred pounds of soybean meal. Thus, when the price of the complementary corn rose, and the price of the substitute, soy meal, fell far enough, feed manufacturers could stop buying manufactured lysine (Connor 2007a, pp. 199-201).
ADM’s main line of defense was to criticize the simple before-and-after analysis (White 1996, 2001). The major flaws in the plaintiffs’ method were alleged to be: (1) The benchmark price would have been generated by noncooperative oligopolistic behavior rather than purely competitive conduct; (2) the price increases observed after the summers of 1992 and 1993 could have been seasonal rather than conspiratorial; and (3) the affected period chosen by the plaintiffs’ expert was too long. As a matter of legal strategy, it is worth noting that because the lysine defendants had not yet admitted guilt, their experts would not be expected to present alternative overcharge estimates to the judge. Thus, trying to demonstrate weaknesses in the plaintiffs’ case was virtually the only option available to obtain a fairness ruling favorable to the defendants.

All three of these criticisms are logical possibilities. All of the economists working on the case agreed that the U.S. (and indeed the global) lysine industry was a classic oligopoly. Sales concentration was high (the Herfindahl index above 3000), buyer concentration was low, the product is perfectly homogeneous, and several barriers to entry were present. "In sum, the lysine industry had virtually all the characteristics of an industry in which implicit oligopolistic coordination of some kind would likely have arisen in the absence of the explicit conspiracy" (White 2001, p. 28). If true, the but-for equilibrium price would, according to most oligopoly theories, be above the competitive price, and the overcharge significantly smaller. However, a few features of the lysine industry may have prevented the sellers from forming an oligopolistic consensus. Chief among them is the fact that two of the five sellers (ADM and Cheil) were brand new to the industry; in other words, the major players did not have a sufficiently long history of strategic interaction to form conjectures likely to yield a stable implicit agreement. Prior to its entry into the lysine industry, ADM had no overlapping markets with any of the three incumbents (Connor 2000, app. E). Moreover, we now know that the principal form of oligopolistic conduct among the incumbents prior to the 1992-1995 cartel was explicit price fixing. Finally, internal contemporaneous documents show that ADM fully intended to put at least one of the incumbents out of business in order to achieve its announced goal of market-share parity with Ajinomoto (Connor 2007a, ch. 8). Predatory conduct by ADM, most likely to be effective in the U.S. market, could have driven the but-for price down to and even below the long-run competitive price in 1992 and 1993. Had the Asian manufacturers not agreed to join an ADM-dominated cartel, ADM might well have continued predatory pricing well beyond June 1992.

Seasonality of demand for lysine was well recognized by the managers of the cartel (Connor 2007a, pp. 201-202). It arises from swine-feeding practices of producers in the temperate zones. With less than six full years of price data for lysine and no price series for comparable feed ingredients,
it is difficult to be precise about how strong seasonal effects are. " Taking the simple average of prices at the seasonal peak to prices at the seasonal trough (about 14 percent) would cause the calculated overcharge to decline by 10 percent from an estimate that ignores seasonality.

In most forensic settings, the affected period and conspiracy period are treated as identical. Both sides acknowledged that changes in transaction prices lagged changes in posted prices, the latter being the conspirators' direct price-fixing tool. In a retrospective evaluation of the cartel, a defendant's expert contends that a seventeen-to-nineteen-month period is the appropriate affected period (White 2001). He bases his position on the "unusual and suspicious" pattern of "uncharacteristic stability" in transaction prices from September or October 1993 to February or March 1995 (see Figure 11-1). The plaintiff's expert, on the other hand, maintains that information revealed by the Chicago trial provides additional support for a forty-two-month affected period (Connor 2001). Without more formal analyses it is not possible to determine the affected period definitively."

The Cost-Based Approach*

This method and econometric modeling (discussed below) were not available to the economic experts doing battle in the 1996 federal civil proceeding, but they do shed light on the accuracy of the before-and-after method. The disparity between the two sides was an order of magnitude. Plaintiffs considering opting out of the federal class had a $150 million overcharge calculation before them, whereas the defendants' settlement offer of $45 million implicitly supposed a $15 million overcharge by the lysine cartel.

During the 1998 criminal trial of three ADM executives for lysine price fixing, prosecutors introduced the confidential production and sales records of ADM's lysine department as exhibits. These internal records (now public documents) provided ADM managers with monthly plant output and several costs (labor, energy, dextrose, other chemicals, overhead...
expenses, transportation, storage, and sales-office expenses) during the five years 1991-1995. Figure 11-2 plots these costs of manufacturing and distribution against monthly physical plant output using regression analysis.

The plot appears to show considerable “scale economies” for levels of output up to 10 or 11 million pounds per month. In fact, the diagram really captures strong learning-by-doing effects because all of the observations below 11 million pounds are drawn from the precartel period (February 1991-June 1992). Abundant testimony and the manufacturing records themselves support the fact that nearly all of the high-cost months were ones with “yield failures” due to contamination of fermentors.” As ADM learned how to sanitize its plant’s fermentation reactors, contamination episodes ceased, and the costs of spoiled-product disposals disappeared.

The most important feature of the average-total-cost curve shown in Figure 11-2 is the portion above 10 or 11 million pounds per month. During the conspiracy, plant output always exceeded 10 million pounds. Statistically, this portion is completely flat. It is true that manufacturing costs were affected by short-run changes in the price of dextrose, which in turn was closely related to the market price of corn. Nevertheless, total manufacturing costs hewed quite closely to the average of $0.63 per pound whenever production exceeded 10 million pounds. As plant output edged closer to the maximum 18 million-pounds level, unit fixed costs dropped a bit.

“In every case when costs jumped above about $0.80 per pound, the lysine/dextrose yield ratio dropped below 30 percent. Such episodes became rare after June 1992 or when monthly production was above 10 or 11 million pounds.”
However, the decline in unit fixed costs was nearly perfectly balanced by higher selling costs that were incurred as ADM shipped higher shares of its U.S. production to overseas destinations. Thus, after June 1992 (the likely cartel period), average total accounting costs of manufacturing and sales varied only within the range of $0.73 to $0.78 per pound and were statistically unrelated to the quantity produced. Adding a fairly generous return on investment of 6 percent of sales brings the average total economic costs to $0.77 to $0.83 per pound of lysine."

Whether in competitive or monopolistic industries, profit-maximizing firms equate their marginal revenues to their marginal costs. Normally, marginal costs are unobservable, but because ADM’s total costs were effectively constant during the cartel period, then average variable costs are equal to marginal costs. It follows that the but-for competitive price would have been just about $0.80 during the affected period. This observation is reinforced by the fact that ADM’s costs of production were equal to or lower than those of all four of its rivals in the lysine industry (Connor 2001, p. 217). The full-cost price of $0.80 seems like a defensible but-for price.

Econometric Modeling

With sufficient time and access to detailed price and cost information, statistical simulation modeling is often the preferred analytical approach of forensic economists in estimating antitrust damages (Slottje 1999). When all the necessary data have been received from defendants during discovery and cross-checked for accuracy and completeness, the econometric analysis itself could take several months longer.

With a rigorous model that is shown to fit the market’s actual performance over time, the legal goal of isolating the effects of a defendants’ illegal conduct from all other market forces is feasible. Econometrics seems ideally suited to identifying "... the only causal factor accounting for the difference between plaintiff’s actual experience in the damage period and its but-for period ..." (Page 1996, p. 36). Law journals and handbooks for lawyers in the antitrust field frequently include material on regression analysis for damages calculations (e.g., Fisher 1980, Page 1996, and Hovenkamp 1999). Baker and Rubinfeld (1999) and Bränder and Ross (2006) provide surveys of the method.

Econometric estimation usually employs reduced-form equations using multiple regression methods." The structural model assumes that

"This is generous because it is ADM’s own rate of return during fiscal 1990-1995 when its profits were bloated by several commodity cartels (Connor 2000, app. A). It is also well above the average return earned by publicly traded companies in similar industries.

"Price effects could be measured from a simultaneous equation system built upon Cournot or Bertrand pricing assumptions. The greatest advantage is the fact that data only from the cartel period would suffice. The main disadvantages are that it is more time consuming, relies on accounting costs
both the quantity demanded and the quantity supplied are dependent simul­
taneously on the market price. One of two common econometric approaches
regresses market price on observations from both the collusive and non-
collusive periods. The right-hand side of the equation contains variables
representing variable input costs (wages, materials, energy, inventories, and
possibly exchange rates) and variables that capture sources of shifting
demand (customer incomes, buyer output levels, seasonal dummies, and
prices of substitutes). The key variable is a qualitative variable taking a
value of one for each period during which the cartel is assumed to have
effectively raised prices. The coefficient of this time-based dummy variable
is the unit markup overcharge. Such models are in essence elaborations of
the before-and-after method, but they are better able to handle multiple
exogenous shifts in demand and supply.

If the analyst believes that price fixing may have influenced the costs of
production, then the dummy variable for time may underestimate the price
effect of collusion. In this case, the appropriate approach is a second method:
fit a reduced-form regression to only a noncartel period. The regression coef­
ficients on all the independent variables are then used to forecast (or back­
cast) the but-for price during the cartel period. In this approach, all the
demand and supply variables can vary during the cartel period. Morse and Hyde (2000) developed and tested an econometric model of
the lysine industry using 1990-1995 monthly information. The model
incorporates a fairly complete list of determinants of lysine demand: the
number of hogs needed by U.S. slaughterhouses, red meat and poultry
export demand, the prices of a complement (corn) and a substitute (soybean
meal), and the seasonality of lysine demand. On the supply side, an equa­
tion related ADM’s U.S. production to the costs of three principal inputs:
dextrose, other variable costs of manufacture, and capital. Both of these
equations fitted the data quite well, and the signs were those predicted by
economic reasoning. An equation permitted them to measure the degree of
competitiveness (“conjectural variations”) between ADM and its four
rivals. Morse and Hyde deduced that the U.S. lysine-cartel overcharge was
about $80 million.

Econometric modeling has become the world standard for proving car­
tel damages.” Nevertheless, econometric estimation has some disadvantages

supplied by the defendants, is more difficult to explain to lay audiences, and is quite demanding with
respect to data on market structure. Such models are unusual in forensic settings.

“Bernheim (2002) makes the interesting point that under some conditions an overcharge estima­
tion in the U.S. market need not adjust for exchange rates even though imports account for the
majority of supply. This is true, for example, when the raw materials used to produce the cartelized
product (e.g., petroleum derivatives) are denominated in U.S. dollars.


“Dynamic simulation techniques may be an alternative for creating a but-for market scenario. The
economist starts with a structural model of oligopoly and calibrates key parameters in the model
using observations from the natural market of interest (Froeb and Werden 1996). If the analyst
compared with the other methods mentioned. Compared with other methods, it is data-hungry: Time series of dozens of demand or cost variables may have to be collected, and less than forty or fifty noncartel observation periods may produce statistically fragile estimates. Specification of the model to be tested and the techniques to correct econometric flaws may become issues too abstruse for the fact-finders to comprehend. Biased estimates may result if postcartel pricing conduct does not return to precartel conduct. Best from a rhetorical stance is to apply statistical methods but supplement them with other approaches and hope that they are mutually supportive.

The Yardstick Method

The yardstick method involves the identification of a market that is similar to the cartelized market but that has prices that were unaffected by the conspiracy. A yardstick market should have cost structures and demand characteristics that are highly comparable with those of the cartelized market yet lie outside the orbit of the cartel’s influence. A geographic yardstick method is applicable to cases of localized price fixing or bid rigging. Markets with nonstorable products, with high transportation costs relative to price, for localized services are good candidates for the yardstick method.” With a global cartel like lysine, the geographic yardstick method could not be applied.

In principle, yardsticks can be developed for analogous product forms sold in the same geographic market as the cartelized product. One might search for a good that is made using a similar major input, that uses a similar technology of production, and that is sold to customers with similar demand characteristics as the cartelized good.” There is, however, makes only a few strong assumptions about the market, the models can become computationally quite complex (Froeb and Werden 2000). “The first major use of simulation in applied industrial organization was, and the most important use still is, in predicting the effects of mergers” (Froeb and Werden 2000, p. 134). Werden (2000) summarizes a simulation of a proposed merger between two leading U.S. bread makers. There do not appear to be any examples of simulation techniques being employed to estimate the size of damages in cartel litigation. However, there are a few academic studies that may point the way to forensic applications. Raper et al. (2000) use simulation to determine the degree of monopsony power exercised by cigarette manufacturers in the U.S. market for leaf tobacco. De Roos (2004) provides an example of how well dynamic simulation can fit the facts of a cartel, in this case the global vitamin C conspiracy.

“The yardstick method has been used in markets for bread (Mueller and Parker 1992), fluid milk (Porter and Zona 1999), and construction services.

Choosing an analogous market requires judgment that is best informed by a deep study of the market and cartel practices. An appropriate yardstick should be justified by examining price movements in the affected market and the analogous market before and after the cartel period. If monthly prices for two or three years are highly correlated, then the analyst has some assurance that the analogous market was a proper choice.

For lysine, methionine was the leading yardstick candidate. It is an amino acid demanded by feed manufacturers and made by fermentation. However, the EU determined that a secret, widespread methionine cartel had operated nearly simultaneously with the lysine cartel. Had methionine price
a logical problem with product-form yardsticks in cartel cases. Cartels can work only if they sell a well-defined product, and such products by definition have no close substitutes. A good product yardstick is likely to be a close substitute for the cartelized product over a wide range in observed prices. Thus, it appears to be a nearly hopeless task to find a good product-form yardstick that has a price unaffected by the alleged conspiracy.

Using Game Theory to Check Results

The defendants in the lysine cartel provided a rebuttal to the plaintiffs’ before-and-after analysis. They asserted that a noncooperative form of collusion was more probable than perfect competition had the cartel not operated (Warren-Boulton 1995). Further, the defendants specified the homogeneous Cournot model as the most appropriate one because of its long-standing acceptance and widespread analytical use in economics. Over certain ranges of market conditions, that model predicted equilibrium prices that fell within the range of actual market prices observed during the cartel period. That is, the Cournot model implied that the cartel had been ineffective in raising prices by explicit collusion above prices that would have been generated by implicit (and legal) pricing coordination. Thus, he asserted, the overcharge was zero.

Predictions from specific oligopoly models require structural parameters. In particular, the Cournot formula for calculating the profit-maximizing price ($P$) needs three pieces of market information: the Herfindahl index of concentration ($H$), the own-price elasticity of demand ($\eta$), and the marginal cost of production ($C$).

"Of course, the "cellophane fallacy" demonstrates that above a certain price, any cartelized product will run into product substitutes. In the fermented-lysine cartel, the cartel’s managers were well aware that when the price of soybean meal was low enough, its natural lysine content would become price competitive with the fermented version."

"Connor and Bolotova (2006, p. 1134) found that overcharge estimates derived from yardsticks were systematically overstated.

"The formula \[ P = C/(1 + \frac{H}{17}) \] is derived from the first-order condition of profit maximization assuming Cournot conduct (Carlton and Perloff 2005, p. 283). Firms are identical, and each firm can supply the entire market.

"Implicitly this assumes that a global cartel was viewing the U.S. market as geographically distinct from others. Internal records of the cartel’s pricing decisions and its efforts to prevent geographic arbitrage tend to support this view. Global concentration was about 2500 in 1994 (Connor 2001a, table 8.A.3)."
(or close to it), and the elasticity was around -0.5 to -1.0 during the cartel period."

One problem with the Cournot model is that the formula can, under some ranges of parameters, predict impossible prices. That is, the model can "blow up." For example, given the high degree of market concentration, if the demand for lysine is highly inelastic (less than \(-0.35\)), then Cournot oligopolists would be predicted to set negative prices, no matter what the cost of production. Negative prices are rarely observed in natural markets because prices generally must be set above the variable costs of production, and these costs are always nonnegative. Another problem with Cournot is that it is only one of many plausible oligopoly models; its popularity with economists rests more on its mathematical tractability than its consistency with the organization of natural markets. Given the lysine parameters just discussed, other equally plausible models such as price leadership by ADM produce equally untenable market price predictions. Finally, although possibly allowable as evidence in antitrust cases, the degree of economic literacy required to comprehend formal oligopoly models greatly restricts their use in forensic settings.

CONCLUSION

One of the hallmarks of a rigorous scientific discipline is the ability to measure parameters of interest with precision. From this perspective, the highly variable estimates often presented of cartel overcharges could be interpreted as reflecting badly on empirical economics. For example, the lysine-cartel overcharge estimates varied by as much as a factor of ten when the first civil suit was being resolved. A more sanguine view is that progressive analyses often show a movement toward greater precision, a movement made possible by additional information and the time to apply more complex analytical methods. At the very least, testimony by opposing experts may yield a range of damage estimates within which parties will settle or a fact-finder will choose a compromise.

Modern cartel enforcement is a paradox. The stated goal of antitrust laws of most nations is deterrence, and optimal deterrence requires that cartel penalties be based on multiples of economic injuries corrected for the probability of punishment. Yet, antitrust authorities are typically reluctant to calculate fines on the basis of damages because of perceived analytical challenges (ICN 2005). However, rough but reasonable estimates can be

Connor (1996) opined that feeds were manufactured under fixed proportions, which implied poultry, swine, or meat elasticities of \(-0.10\) to \(-0.50\). These are retail-level elasticities calculated from precartel, more competitive periods; at the higher cartel-period prices, the elasticity will be higher in absolute value. This discussion took place in the context of his analysis of the deadweight loss from cartel pricing.
quickly prepared using one of several methods delineated in this chapter, especially when the analyst has sufficient appropriate economic data. More often than not, alternative estimates of cartel overcharges tend to be mutu­ally supportive

In July 1996, the trial judge approved the $45 million offer as fair and reasonable. His decision seems to have been made for reasons unrelated to the economic evidence presented. Rather, he seemed most persuaded by the testimony of class counsel that this was a hard-fought deal that was unlikely to be improved. About thirty-three of four hundred direct buyers in the class begged to differ by opting out of the settlement. Most of the opt-outs were larger firms with the legal resources to continue hard negotiations with the defendants and with the patience to wait for their rewards. Although settlement terms are confidential, reports in the media suggested that the opt-out firms, with the benefit of criminal guilty pleas by the lysine-cartel members, received at least double the amount per dollar of purchases received by the smaller buyers in the class (Connor 2007a, pp. 397-398). This trend—opt-outs becoming a larger share of the buyers and obtaining better settlements—continued in subsequent cartel settlements throughout the late 1990s (Connor 2007a, pp. 402, 406).

With the benefit of hindsight and a great deal more information, it appears now that the initial $150 million estimate by the plaintiffs was too high. Considering seasonality would have reduced the overcharge amount by about 10 percent or so. More importantly, the two periods selected to determine the but-for price were most likely unrepresentative predatory-type episodes that could not have been sustained for the entire three years of the cartel’s operation. That is, ADM was punishing its future cartel partners by unilaterally forcing prices down to the point where ADM was barely covering its variable costs. A $.70 but-for price was not enough to cover full economic costs. In retrospect, a but-for price of $.80 was closer to a long-run competitive price, suggesting that the true U.S. overcharge was around $80 million.

The federal lysine class and the opt-outs from the class eventually collected approximately $70 million from the cartel members; indirect purchasers of lysine obtained about $25 to $30 million in state courts where such buyers have standing to sue. Thus, U.S. lysine buyers recovered as a group slightly more than single damages; net of legal fees, buyers recovered less than single damages (Connor 2007a, table 18.1). The lysine buyers’ settlements were comparable with the settlements in two similar cartel cases that followed on the heels of the lysine case (Connor 2007a, p. 474). Direct buyers of citric acid and vitamins recovered 90 percent and 135 percent, respectively, of U.S. overcharges. These results reinforce Lande’s (1993) conclusion that civil price-fixing awards are typically less than actual damages (and considerably below treble damages).

The ex ante perception of antitrust liability is a critical determinant of deterrence. If would-be price fixers expect that their monopoly profits will
exceed the financial costs of antitrust fines and civil settlements, it is rational for them to form a cartel. In the "Introduction" section above, it was noted that the potential antitrust liability in the United States is currently approximately seven times the U.S. overcharges, and overcharges are only slightly larger than the monopoly profits from collusion. One might argue that the best conjectures that a company could have made about antitrust liability during 1988-1992 were much lower because of the changing legal landscape during the latter part of the 1990s. However, this is at best only a partial explanation for the spate of global cartels unmasked and punished after the lysine actions in 1996. In evaluating the deterrence capabilities of current antitrust sanctions, one must also consider three additional factors: the probability of an operating cartel being detected and prosecuted, actual rather than maximum antitrust penalties, and the geographic location of the monopoly profits generated by a cartel.

Limited evidence from the United States and Europe suggests that the probability that an established illegal cartel will be caught is somewhere between 10 percent and 30 percent (Connor 2007b, table 1). Outside of North America and western Europe the chances are negligible. Evidence from the three best-documented cartel cases of the late 1990s shows that U.S. criminal and civil penalties approached but never exceeded double the U.S. overcharges. In the EU, these same cartels were made to pay fines about equal to single damages, and the prospect of civil damages suits in Europe is dim. In Asia, cartel fines are small or nonexistent. A striking feature of most global cartels discovered since 1996 is that their sales and profits were distributed almost equally across North America, western Europe, and the rest of the world. Taking into account each of these three general features of modern cartels and contemporary anticartel enforcement practices, it logically follows that global-cartel deterrence requires actual monetary sanctions to exceed three to ten times worldwide overcharges, where the exact multiplier depends inversely on the probability of discovery. Treble damages in the United States alone will not deter global cartels.

The rationale of deterrence can be illustrated with financial information about ADM, the most heavily fined of the lysine conspirators. ADM's monopoly profits from fixing the U.S. price of lysine for three years were about $80 million. ADM paid a $70 million fine to the federal government, about $49 million to direct buyers of lysine, and $15 million to indirect buyers in state court cases. Not counting legal fees and other intangible costs, ADM's ex post U.S. costs of collusion somewhat exceeded its U.S. revenues from collusion. On the other hand, ADM garnered approximately $100 million from its non-U.S. sales of cartelized lysine. Non-U.S. fines left ADM with a positive return of $25 to $35 million on its non-U.S. price fixing.

Crime did not pay for ADM in the United States, but it did pay abroad. More importantly, when one factors in the expectation that the chances of being caught were small, ADM's decision to form the lysine cartel was eminently rational.
REFERENCES

ADM. Annual SEC Form 10-K Report for the Year Ending June 30, 2001 and previous years.
Connor, John M. Lysine Production, Trade, and the Effects of International Price Fixing, 3d edn. West Lafayette, IN: Department of Agricultural Economics, Purdue University, 1999.


Morse, B. Adair, and Jeffrey Hyde. Estimation of Cartel Overcharges: The Case of Archer Daniels Midland and the Market for Lysine, Staff Paper 08-00. West Lafayette, IN: Department of Agricultural Economics, Purdue University, October 2000.


INTRODUCTION

Sometime between May and August every year, school district officials throughout the country independently solicit bids on annual supply contracts for milk and other products. In response to the solicitations, dairies that are in a position to supply school milk submit bids for procurement contracts. The majority of these procurements employ sealed-bid auctions, in part because federal guidelines require the purchasing method to be as competitive as possible. Typically, the low bidder is selected to supply milk in half pints to the schools during the following school year. According to a study by the U.S. Department of Agriculture, 86 percent of dairy purchases for the 1996-1997 school year used a formalized bidding procedure. The value of the contracts totaled more than $700 million.

In 1993, representatives of two dairies from Cincinnati, Ohio (Meyer Dairy and Coors Dairy), confessed to rigging bids in school milk auctions in the 1980s. These individuals testified that they had rigged bids with each other and with Trauth Dairy. Their testimony was offered as part of a settlement of the criminal and civil cases against them. Together they paid several million dollars to settle the cases. The bid-rigging scheme described in their testimony was one of respecting incumbencies. If one of the cartel members had served a particular school district in the previous year, then...
other conspirators were to submit high complementary bids that would not undercut the incumbent firm’s bid, or else refrain from bidding. The testimony described frequent communication among these competitors as the specific details of the scheme were worked out through the bidding season.

Trauth Dairy, another Cincinnati dairy that was implicated by the whistle-blowers, maintained that there was no conspiracy. Instead, it claimed that the statements by its competitors were designed to ruin it and drive it from these markets, to the benefit of the whistle-blowers. It maintained its innocence through several trials, criminal and civil.

In 1994, the state of Ohio charged thirteen dairies, including Meyer, Coors, and Trauth, with collusion and bid rigging in school milk auctions for the years 1980 through 1990 inclusive. Collusion in auctions is an agreement among a group of firms that is designed to limit competition between the participants. The analysis of collusion or communication among bidders is more difficult than the analysis of competitive bidding, in part because the nature and effects of conspiracy depend on the specifics of the situation.

Because the direct evidence of the extent of the conspiracy was disputed, economic evidence was gathered in addition to other kinds of evidence as the state set out to make its case. Three types of economic evidence were collected: (1) evidence on the incentives to collude in these markets; (2) evidence as to whether the behavior of the alleged conspirators was more consistent with competition or collusion; and (3) evidence on the extent of damages caused by the alleged bid rigging.

The first category of economic evidence concerns the incentive to collude. As we describe below, the details of the procurement process, the nature of milk processing and delivery, and the characteristics of demand for school milk are such that collusive agreements among suppliers may be relatively easy to reach and maintain. Collusion appears to be a pervasive phenomenon in school milk auctions. There have been investigations of price fixing in auctions for the provision of school milk in more than twenty states, beginning with a Florida bid-rigging case in 1988. More than 130 federal criminal cases have been filed, and many states have also filed charges. Guilty pleas have been entered in at least a dozen states, and fines have been levied in excess of $90 million. About ninety people have been sent to jail for average sentences of six months.


See, for example, Henriques and Baquet (1993) and Lanzillotti (1996).
We then analyze the behavior of the defendants relative to a competitive benchmark. We conclude that the bidding behavior of the accused dairies was more consistent with collusion than with competition. For example, some of the defendants exhibit patterns of both local and distant bid submissions. That is, they submit bids relatively near their plants and they also submit bids well beyond their local territories. Further, the distant bids by the defendants tend to be relatively low. In contrast, bids by a group of firms that were not named as defendants are an increasing function of the distance from the school district to a firm's nearest plant. These features of bidding by the defendants are consistent with territorial allocation of districts close to the dairies' plants to restrict competition and relatively competitive bidding at more distant locations, which were perhaps outside the area of territorial allocation. If bidding for local districts had been competitive, local bids should have been lower than distant bids, because shipping costs were lower and because the Cincinnati area has several potential local suppliers. The relationship between bidding behavior and distance is notable, because processed milk is relatively expensive to ship (since its value is low relative to its weight) and competition is therefore localized.

Third, we also analyze the effect of the collusion on prices paid by the school districts in Ohio as a result of the collusion as identified. We estimate that collusion caused school districts to pay about 6.5 percent more for school milk in the affected area during the period analyzed, and much more than that in some locations.

CHARACTERISTICS OF OHIO SCHOOL MILK MARKETS

Market outcomes are determined by three factors: the nature of demand, the nature of the production process, and the nature of competitive interaction among suppliers. In the case of school milk, demand and cost characteristics are relatively easily described. The more difficult problem is determining the nature of competitive interaction given the demand and cost characteristics; that is the focus of a subsequent section.

There are more than six hundred school districts in Ohio. Most districts award annual contracts for the supply of school milk. For example, a district will indicate that it expects to purchase fifty thousand half pints of whole white and thirty thousand half pints of whole chocolate milk. (This corresponds to a student population of about 450.) A bid submission consists of a list of prices for the various products. In addition to specifying categories of milk, the district may require its supplier to provide coolers (for the refrigeration of milk), straws, or napkins. Escalator clauses with price indexed to the price of raw milk are also provided in some bid requests to reduce the risk to the dairy that is associated with submitting a bid on an annual contract, since the price of raw milk can fluctuate substantially over a year.
Demand for school milk is relatively insensitive to the price charged. School milk programs are subsidized, and milk demand may be inelastic in any event. An estimate of the responsiveness of school districts’ demand for milk to price, using data on a cross-section of school districts in Ohio, indicates that price is not a statistically significant determinant of demand despite substantial variations in price. If school district purchasing is inelastic, a firm that controls the supply of milk for a particular district could profit from a substantially higher price in that district.

The competitive price will depend on the costs of potential suppliers. We now describe the most important components of these costs. Dairy processors receive raw milk from dairy farmers in the area. The price charged for raw milk is typically regulated through an elaborate federal marketing order (FMO) system based on location and type of milk. Although some milk producers are unregulated, the market would tend to produce a single price for both regulated and unregulated raw milk in the long run. The cost of the raw milk contained in a typical half pint purchased by a school district is about seven cents. (This figure varies seasonally and from year to year.) Raw milk is processed by standardizing the butter fat content (e.g., 2 percent or skim) and then pasteurizing, packaging, and delivering it. Typically, potential suppliers of school milk pay about the same amount for raw milk, and they use much the same technology to pasteurize, package, and deliver. Many firms use the same suppliers of packaging materials, typically paying about two cents per half pint for packaging. As described below, we would expect incremental costs to be similar across different suppliers or potential suppliers. Any differences in long-run incremental costs would likely arise from differences in distance from the plant to the district.

When bids are solicited by school districts, firms in a position to supply will submit a bid. Which firms are in a position to supply the district? A firm must have access to a supply of school milk. The potential suppliers of school milk fall into one of two categories: processors, which process and package raw milk in the half pint containers demanded by schools; and distributors, which often purchase milk wholesale from the processors and resell it to a school district. For any firm interested in selling its own milk in school milk markets, the costs of a processing plant may represent a substantial entry barrier. For those processors for which we have data, the school milk business represents a small fraction of their total revenues, typically less than 10 percent. Although a processing plant is necessary to process

*There are many distributors active in Ohio school milk markets, but the relationship between processors and distributors can be classified into relatively few categories. First, some distributors are granted exclusive territories and offer bids using the affiliated processor’s name or stationery. Second, some distributors purchase milk wholesale from one or more processors and sell it to schools using their own name. These distributors often switch processor affiliations from year to year, presumably to get better terms. Finally, but not as frequently, distributors can provide delivery to the processor for a fee. Since all three arrangements coexist in the market, any particular type of distributor/processor relationship is unlikely to convey a competitive advantage.
school milk, school milk is only one of many products. To our knowledge, no firm has ever built a processing plant solely to supply school milk.

Thus, the decision to build a plant hinges on considerations in other product markets, such as supplying wholesale to supermarkets or to other institutions such as restaurants. (It should also be noted that there was substantially more exit than entry of processing plants in Ohio during the 1980s.) The costs of the processing plant that are directly attributable to school milk are therefore quite small. We conclude that only those firms that have access to a milk processing plant, largely put into operation for other reasons, would have the ability to enter the school milk business selling their own product.

If a firm has a plant that is relatively close to the school district in question, what costs would be incurred should the firm submit a bid and win the contract? For example, are there any important fixed costs associated with serving school districts? Because data on the fixed costs of supplying school milk are not available, we analyze the size distribution of firms in Ohio. For those companies participating in school milk markets, the scale of operations ranged from one firm supplying milk to about 1 percent of Ohio public school students annually to another that supplied about 7 percent of the students. Firms producing on vastly different scales coexist in these markets, even though marginal costs seem to be quite similar, indicating that the fixed costs associated with the school milk business for a dairy serving other customers are likely to be relatively small.

Given that a firm has a processing plant and is in the school milk business, there are incremental costs associated with starting service to a new school district, such as the costs associated with providing a cooler and adding the district to existing delivery routes. The delivery portion of total incremental cost may also be related to the number of deliveries per week for the new district and how well it fits with the firm's existing route structure. There are three main options for delivery: some dairies have dedicated school milk routes; others add school deliveries to their regular retail commercial routes; and other dairies use local delivery subcontractors for these services. Frequently, a dairy will utilize more than one option. Of course, delivery costs vary with distance. A typical delivery cost is on the order of a penny per half pint for local deliveries.

The cost of coolers is small relative to the other costs of school milk. For example, a sixteen-case cooler can be purchased currently for $1100. Using an eight-year depreciation schedule, which probably underestimates a cooler’s useful life, this figure implies a cost of about 0.2 cents per half pint for a cooler that chills approximately seventy thousand half pints per year.

As noted earlier, firms can and do enter the market for the distribution of other firms’ processed milk. Thus, entry barriers into distribution would be lower.

Absent collusion, bidders will not know exactly which school districts they will ultimately serve when they bid on individual contracts. That is, they will not know the outcome of future contract lettings.
In summary, the cost of ingredients for a half pint of milk is on the order of about seven cents, packaging costs are about two cents, and delivery costs to serve a nearby school district are about one cent. Total delivered incremental costs are on the order of ten cents per half pint during the period analyzed. We expect these incremental costs to be independent of the scale of operations and similar across firms (after accounting for proximity). Fixed and other one-time costs will affect a dairy’s decision whether to enter or exit the school milk business, whereas the incremental costs of supplying an additional half pint of milk or servicing another district are relevant for its pricing decisions.

The market characteristics described so far suggest competition among firms producing a homogeneous product with similar and constant incremental costs. Firms are likely to have good information about the costs of their competitors, since most cost changes affect all firms similarly. There are unlikely to be substantial informational advantages in the market. School districts would be willing to pay high prices for school milk, if they had no other choice. Competition is likely to be localized, because of the regulation of raw milk prices and because of relatively high transportation costs.

FACTORS FACILITATING COLLUSION

There are many features of Ohio’s school milk markets that may affect the dairies’ competitive interaction. Specifically, there are a number of characteristics that facilitate collusion:

1. Firms compete only on price. Under the terms of the contract, the winning bidder supplies the product with specified characteristics (e.g., butter fat content or flavoring). A cartel need only coordinate submission and bid decisions, and not other characteristics of the product, which simplifies cartel operations.

2. The school districts’ policy of publicly announcing bids and the identity of the bidders allows cartel members to detect “cheating” from cartel agreements. Undercutting and cheating on collusive arrangements would not go unnoticed, so a collusive arrangement is therefore more likely to be stable.

3. Most school districts held their auctions annually but at different times during a year, and they acted on their own. The disorganized letting of contracts (as opposed to all contracts being let on a particular day, for

1These characteristics are prevalent in other markets. See, for example, Porter and Zona (1993).

2Some school districts occasionally band together in cooperatives and solicit bids for the group. However, cooperative arrangements historically have not prevented districts from soliciting bids individually.
example) allows cartel members to adjust bidding behavior during the bidding season to allocate market shares, and it provides an opportunity for nearly immediate retaliation for "bad cartel citizenship."

4. The predictability of the demand for school milk from year to year allows threats of future retaliation in response to deviations to be credible.

5. The fact that the markets themselves are easily defined according to school district boundaries permits allocations by assignment of territories.

6. The set of firms potentially submitting bids in a particular market is small and quite stable.

7. These firms use similar production processes and therefore face similar cost structures. The similarity of potential suppliers makes it more likely that a group of firms could agree on joint behavior.

8. The same dairies encounter one another in more than one market, and so competition may not be fierce. Contact between competitors in multiple markets (school districts as well as wholesale accounts) makes collusive schemes that allocate markets more feasible.

9. The practice of obtaining competitors' price lists through retail customers (e.g., grocery stores) was common, if not universal. Advance notification of list price increases can lead to supracOMPETITIVE prices: The practice allows communication of intentions of competitors (Holt and Scheffman, 1987).

10. Dairies are frequently customers of one another. This facilitates direct communication and allows a pretext for meetings between competitors. The practice allows communication of pricing information (even for products that are not purchased) through full price lists. Holt and Scheffman (1988) condemn this practice as potentially facilitating collusion.

11. The many dairy trade associations in Ohio also allow a pretext for meetings of competitors. Most meet on a regular basis to discuss issues of mutual interest. These associations obviously understand the legal dangers since the minutes of these meetings often indicate that they begin by reading a statement warning members not to discuss prices.

According to the testimony of the Meyer and Coors representatives, these factors were of more than theoretical importance. Even if one ignores their testimony, a cartel among firms operating in the area is plausible and

Bernheim and Whinston (1990) show that multemarket contact in a repeated market setting "relaxes the incentive constraints that limit the extent of collusion."
would be to the benefit of each of the participants. Since competition is localized, prices will fall to competitive levels only in areas where there is a sufficient number of local competitors. Distant competitors are disadvantaged by transportation costs and can discipline price increases to only a certain extent. The competitive significance of each supplier is directly related to its relative distance from the school district. Any elevation of price that is achieved through collusion could, in some circumstances, be defeated by the entry of new firms from outside the market. In the case of school milk, entry will not take the form of a new firm’s setting up a processing plant; instead, it will be in the form of bids from firms whose plants are farther away. In this way, the transportation cost for distant noncartel entrants will constrain the ability of a cartel to raise prices, but if collusion is effective, then prices could be elevated by a significant amount. If local firms could coordinate their bidding behavior, then some profits could be earned without stimulating entry from distant suppliers. The reaction of any school district is unlikely to limit the ability of a cartel to raise price. As noted above, demand for school milk is inelastic, and school districts would continue to purchase milk even at elevated prices, to the detriment of the school districts and to the benefit of cartel members.

A scheme of respecting incumbencies is one way for a cartel to coordinate bidding behavior. Such a scheme would be attractive if cartel firms were near each other, as some are in this case. This mechanism avoids the cartel problem of allocating school districts of varying profitability (due to a location or level of service required) among the cartel members. Another way of coordinating bidding behavior would be the assignment of geographic territories to individual firms. Territorial assignment would be a more practical collusive mechanism when firms are geographically separated and some firm has a clear advantage in serving a particular district.

A MODEL OF COMPETITIVE BEHAVIOR

In a competitive market, firms in the school milk business face two interrelated decisions. First, should the firm submit a bid in a particular district? Second, if a bid is submitted, what should the bid level be? We address these decisions in turn in the subsections below. Here we describe a model of competitive bidding.

Porter and Zona (1999) estimated an econometric model of bidding behavior using a data set provided by the state of Ohio that contains information on school milk procurement for as many as 509 of the approximately 600 school districts in the state from 1980 through 1990, inclusive, with approximately sixty different bidders participating at some point in the sample. We created a control group composed of nondefendant firms that bid on Ohio school milk contracts. The data set contains information on the identity of the districts, their location and enrollment, and the timing of

Page 336
contract lettings; and information about who submitted bids, and the nature of winning and losing bids, such as prices and compliance with district specifications, that were submitted to each district. Table 12-1 provides some descriptive statistics for each year in our sample.

According to data obtained from the U.S. Department of Agriculture, a total of sixty-eight milk processing plants had sales in Ohio at some point during the sample period and potentially could have supplied school milk. Figure 12-1 displays the location and owners of the forty-six plants that supplied school milk in 1987. Superimposed on the plant location map in Figure 12-1 is an index of dairy concentration. The darker the area, the more concentrated are the ownership and control of local school milk processing facilities. Figure 12-1 shows that the markets in the northeast section of Ohio are the least concentrated, and if they are competitive they should have the lowest prices, all else equal. Columbus and the southernmost tip of Ohio have markets where the supply of school milk is concentrated in the hands of relatively few producers. Over the period 1980-1990, the number of plants serving Ohio school districts fell from fifty-four to forty-three, and the number of firms fell from forty-three to twenty-six.

Table 12-2 shows the distribution of bidder distances in the Ohio data. The table also shows the distribution of potential supplier distances in the

---

**TABLE 12-1**

**Characteristics of Ohio School Milk Data Set**

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of School Districts or Cooperatives</th>
<th>Total Enrollment in School Districts</th>
<th>Number of Plants Operated by Submitting Dairies</th>
<th>Number of Dairies Submitting Bids</th>
<th>Average Price per Half Pint ($0.00)</th>
<th>Average FMO Price for Raw Milk</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980</td>
<td>366</td>
<td>1,257,925</td>
<td>54</td>
<td>43</td>
<td>0.1282</td>
<td>0.0718</td>
</tr>
<tr>
<td>1981</td>
<td>398</td>
<td>1,379,619</td>
<td>50</td>
<td>44</td>
<td>0.1295</td>
<td>0.0769</td>
</tr>
<tr>
<td>1982</td>
<td>415</td>
<td>1,371,164</td>
<td>46</td>
<td>40</td>
<td>0.1295</td>
<td>0.0760</td>
</tr>
<tr>
<td>1983</td>
<td>436</td>
<td>1,415,281</td>
<td>48</td>
<td>40</td>
<td>0.1288</td>
<td>0.0764</td>
</tr>
<tr>
<td>1984</td>
<td>448</td>
<td>1,430,644</td>
<td>53</td>
<td>38</td>
<td>0.1313</td>
<td>0.0741</td>
</tr>
<tr>
<td>1985</td>
<td>463</td>
<td>1,460,697</td>
<td>48</td>
<td>35</td>
<td>0.1322</td>
<td>0.0708</td>
</tr>
<tr>
<td>1986</td>
<td>481</td>
<td>1,457,437</td>
<td>47</td>
<td>36</td>
<td>0.1304</td>
<td>0.0700</td>
</tr>
<tr>
<td>1987</td>
<td>494</td>
<td>1,566,591</td>
<td>46</td>
<td>33</td>
<td>0.1310</td>
<td>0.0701</td>
</tr>
<tr>
<td>1988</td>
<td>509</td>
<td>1,520,635</td>
<td>44</td>
<td>30</td>
<td>0.1338</td>
<td>0.0666</td>
</tr>
<tr>
<td>1989</td>
<td>491</td>
<td>1,564,869</td>
<td>43</td>
<td>26</td>
<td>0.1389</td>
<td>0.0708</td>
</tr>
<tr>
<td>1990</td>
<td>412</td>
<td>1,296,587</td>
<td>43</td>
<td>26</td>
<td>0.1575</td>
<td>0.0797</td>
</tr>
</tbody>
</table>

Notes: Dollars per hundredweight FMO (federal milk order) #33 Class 1 Fluid Milk; 186 half pints per hundredweight; prices quoted for July of each year.

---

"We measure concentration using the Herfindahl index. For market shares we use the fraction of plants within seventy-five miles of the region that are owned by each firm."
data. Approximately 85 percent of the firms supplying school milk in a given year do so from plants located less than seventy-five miles from the school district. Firms considering entry in order to supply milk to distant school districts appear to be disadvantaged by distance, even if they have a plant in operation.

The distribution of processing plants in the state and the apparent disadvantage of shipping long distances cause districts to receive a relatively small number of bids on school milk. For the districts in the data set, 45 percent received only one bid, 34 percent two bids, and 18 percent three bids. Even though there are a large number of potential bidders for any particular auction, very few actually bid. The mean number of bids is 1.8. The small number of bids submitted indicates that dairies may not have important private information. If bidders knew their own costs as well as the costs of other potential suppliers, then it seems plausible that bidding by only one or two dairies would be observed. The low-cost supplier would submit a bid just below the cost of the next-lowest-cost supplier, and the latter would
be indifferent between bidding at its own cost and not bidding. About 43 percent of all bids are submitted by the firms with the plants closest to the district, and 8 percent by the firms with the second-closest plant. Among winning bids, 49 percent are submitted by the firms with the closest plants, and 8 percent by those with the second-closest plant.

### Bid Submission

In a competitive market a firm will submit a bid in a particular district whenever the probability of winning is relatively high and when the expected return covers both the costs of preparing the bid and the incremental costs of supplying the district. A model of bidding behavior should account for variables that reflect the potential bidder’s absolute and relative advantage in serving a district. For example, variables that may be important in characterizing these advantages include whether or not the firm (1) has significant transportation costs to the particular district, as reflected by the distance from the district to the plant; (2) is a distributor or a processor of milk; (3) is the closest potential supplier (and so the most likely low-cost supplier for that district); (4) is the second closest, and hence likely the second-least costly potential provider; (5) is bidding on a large or small school district, as larger districts may require more time and energy to prepare a bid; and (6) can efficiently provide the specified milk under the terms

### Table 12-2

**Probability of Bidding and Winning Conditional on Distance from Plant to School District**

<table>
<thead>
<tr>
<th>Distance in Miles</th>
<th>Number of Districts</th>
<th>Probability of Bidding</th>
<th>Proportion of All Bids</th>
<th>Probability of Winning Bids</th>
<th>Proportion of Winning Bids</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-10</td>
<td>2115</td>
<td>(b) 19.5%</td>
<td>(c) 20.1%</td>
<td>(d) 13.6%</td>
<td>(e) 22.9%</td>
</tr>
<tr>
<td>10-20</td>
<td>3197</td>
<td>14.0</td>
<td>21.7</td>
<td>8.9</td>
<td>4.2</td>
</tr>
<tr>
<td>20-30</td>
<td>3840</td>
<td>7.6</td>
<td>14.2</td>
<td>4.9</td>
<td>15.0</td>
</tr>
<tr>
<td>30-40</td>
<td>4526</td>
<td>5.5</td>
<td>12.1</td>
<td>3.4</td>
<td>12.1</td>
</tr>
<tr>
<td>40-50</td>
<td>5637</td>
<td>2.3</td>
<td>6.3</td>
<td>0.9</td>
<td>4.2</td>
</tr>
<tr>
<td>50-60</td>
<td>6440</td>
<td>1.9</td>
<td>5.9</td>
<td>1.0</td>
<td>5.0</td>
</tr>
<tr>
<td>60-70</td>
<td>5314</td>
<td>1.4</td>
<td>3.7</td>
<td>0.5</td>
<td>2.0</td>
</tr>
<tr>
<td>70-80</td>
<td>6732</td>
<td>1.4</td>
<td>4.7</td>
<td>0.8</td>
<td>4.4</td>
</tr>
<tr>
<td>80-90</td>
<td>5290</td>
<td>1.0</td>
<td>2.5</td>
<td>0.6</td>
<td>2.5</td>
</tr>
<tr>
<td>90-100</td>
<td>4885</td>
<td>1.2</td>
<td>2.8</td>
<td>0.7</td>
<td>2.8</td>
</tr>
<tr>
<td>100-150</td>
<td>26079</td>
<td>0.5</td>
<td>6.1</td>
<td>0.3</td>
<td>6.6</td>
</tr>
</tbody>
</table>

Notes: Columns (b) and (d) report, for each distance category, the percentage of districts in which firms submitted a bid or the winning bid, respectively. Columns (c) and (e) total 100 percent and report the fraction of all bids or winning bids, respectively, in each distance category. Results are based on a total of 2053 submitted bids from control group firms out of 73,965 bid opportunities; 1260 of these were winning bids.
of the contract (whether or not coolers are required, or whether fixed price or indexed contracts are specified, for example).

From available bidding data, we estimated how these factors affect the decision whether to submit a bid by the control group of nondefendants. See Porter and Zona (1999) for more detail. The results from this estimation are then used as a benchmark for comparison with the behavior of the defendants.

The econometric results can be summarized as follows. Processors are more likely to submit bids than distributors, all else equal. This may reflect the fact that distributors tend to run a single school milk route, while processors tend to run several routes with school delivery. Firms, especially distributors, are more likely to submit a bid in one particular direction as opposed to all around their plant. This may reflect the effects of existing route structures. Firms are less likely to submit bids to school districts that are further away from their plants than they are to closer districts. This probably reflects absolute cost differences. Districts that request coolers or straws receive fewer bids than districts that do not. The other specification items do not affect bidding behavior in a statistically significant way.

The likelihood of submitting a bid is a decreasing function of distance. Figure 12-2 displays the impact of distance on the probability of bidding for a hypothetical firm in the control group. To construct this figure we assume that the firm is the closest supplier to districts less than ten miles from its

FIGURE 12-2 Predicted Probability of Submitting a Bid by Distance

Note: Variables that are not related to distance are set to their mean value, except processor is set to one, closest is set to one at distances less than ten miles, and closer is set to one at distances between ten and twenty miles.
plant and the second-closest potential supplier at distances between ten and
twenty miles. The three curves correspond to three different specifications.
There are two steps in the predicted probability of bidding where the firm
loses its locational advantage and becomes neither the closest nor the sec­
ond-closest potential supplier. The probability of bidding by this hypothet­
ic firm is above 50 percent at zero distance, but nearly zero beyond
seventy-five miles.

Bid Level Contingent on Submission

In competitive markets, a firm submitting a bid chooses its bid to maximize
expected profit. A bid cannot be increased indefinitely; the higher the bid,
the lower the likelihood of winning. To the extent that bids are costly to sub­mit, it would not be profit maximizing to submit bids that have no probabil­ity of winning. A firm maximizing expected profit trades off higher profits
against a lower probability of winning. Equivalently, bidders choose their markup over costs. The markup is
affected by the probability of winning the auction, which depends on the
likelihood of other firms submitting bids. The markup is chosen to maxi­mize expected profit given the level of cost. There are two categories of
variables that would tend to explain bidding behavior. First, there are vari­ables that may contribute to cost, such as distance from the plant to the dis­trict; whether or not the bidder is a processor; whether or not coolers or
other items are required or supplied; the number of required deliveries; and
whether or not there is an escalator clause. Second, there are variables that
reflect competitive characteristics of the market: the bidder’s relative cost
advantage and the number and costs of other suppliers. For example, if a
bidder is the closest vendor, then that bidder can bid more than it would oth­erwise without decreasing the probability of winning. A second example is
a variable that affects the probability that a firm submits a bid. Firms that
are likely to submit bids on particular districts are those that have a high
probability of winning and those expecting to earn relatively high profits in
the event that they win the auction.

To analyze bid prices, we constructed a summary measure of each
firm’s bids for the various milk products, with 2 percent chocolate used as
the base product. We measured the effects of the competitive factors
described above on the bid levels of the control group of nondefendants.
Figure 12-3 summarizes how control group bids vary with distance. We plot
predicted bids for a hypothetical control group firm holding variables other
than distance constant. The figure incorporates the effect of distance on
the probability of submitting a bid and the resulting impact on the bid.
Figure 12-3 presents the results for three different specifications of the

control group bidding model. A bid one hundred miles away would be between one and two cents higher than in a district adjacent to the plant, all else equal. Firms in the control group are unlikely to submit bids to districts at these distances from the plant, however. Firms closest to the district have some competitive advantage, but that advantage diminishes with distance.

Bidding by nondefendants is consistent with competitive bidding under standard models of spatial competition, where each firm may exercise local monopoly power. As described above, competition is localized, because of the regulation of raw milk prices and because of relatively high transportation costs.

BEHAVIOR OF DEFENDANTS

Comparison with Control Group Behavior

We now examine the bidding practices of the three dairies located in Cincinnati—Meyer, Coors, and Trauth—and compare them with the control group. We consider both the bid submission decision and the level of the bid contingent on submission.
Porter and Zona (1999) tested for differences between the estimated control group bid submission model and that estimated for each of the Cincinnati dairies. For each dairy, they reject the hypotheses that the defendants submitted bids according to the control group model at conventional significance levels under all the specifications considered. There are also significant differences in the statistical process generating the level of bids. That is, the bidding behavior of each of the Cincinnati dairies differs from the control group. Behavioral differences are not necessarily the result of anticompetitive behavior. We are interested in how the behavior of each of the defendants differs from the control group, given that they differ.

**Comparison with a Collusive Strategy**

We are faced with a standard problem in antitrust economics: distinguishing between competitive and collusive behavior. Porter and Zona (1993) identified collusion in highway construction auctions by (1) focusing on bid levels rather than submission decisions, because conspirators apparently submitted complementary bids; (2) identifying differences in the determinants of cartel bids relative to that of other firms; and (3) observing that cartel bids seemed not to be cost based, except for the lowest cartel bid, in contrast to the bids submitted by other firms. Because these differences exist, we conclude that conspiracy with complementary bidding is more likely than not.

In a study of a nineteenth-century railway cartel, Porter (1983) proposed statistical tests to identify whether competition or collusion is more consistent with observed data based on pricing patterns over time. Some observed price fluctuations do not appear to be the result of demand shifts or changes in observable cost factors. Instead, the observed pattern of occasional price wars, following periods of unusually turbulent market shares, is unlikely to be observed under competition. Under a specific theory of conspiracy, such a pattern is possible. The existence of such a pattern informs an inference of collusion.

Our strategy for the problem at hand is similar. The nondefendant firms behave, on average, in a manner consistent with competition. We have concluded that the Cincinnati dairies behave in a statistically significantly different manner relative to the nondefendant firms. Is it likely that these differences are attributable to idiosyncratic effects of cost or competition? If not, are the differences attributable to independent factors, or are there suspicious patterns of correlation? We now address these two questions.

*Baker and Hresnahan (1992), Harrington (2008), and Porter (2005) discuss related methods of detecting the exercise of market power, or the presence of collusion.*
There are many alternative methods of colluding in auction markets.\(^1\) For example, conspirators could refrain from bidding against each other by allocating exclusive territories. Alternatively, they could submit several bids at inflated levels, where the number of bids may be intended to create the appearance of competition. In either event, the members of the ring know that competition has been limited in the affected markets. Any of these firms know that if they submit a bid they do not have to worry about being undercut by another ring member. Observed bids will differ from competitive bidding because the conspirators have coordinated their actions. The expected winning bid will be higher because conspirators have coordinated their actions, whether or not a conspirator wins the auction.

Since distance is an important factor in the control group model, we focus on that dimension. Moderate increases in shipping distance are associated with large declines in the probability of submitting a bid (Figure 12-2). Similar increases in distance are associated with increases in submitted bids in the control group (about 10 percent at seventy miles in Figure 12-3). We examine the deviations of defendant firms' bidding behavior from the control group predictions in this context.

Consider first the differences between predicted and actual bid submission behavior for each of the three Cincinnati dairies, at various distances. There are some notable patterns. First, all three dairies bid more frequently than the control group model predicts for districts within 30 miles of their plant. Second, Meyer is unusually likely to bid in districts 100 to 110 miles away, and Trauth in districts 60 to 80 miles away.

We also compare actual bids\(^2\) with the control group prediction. The bids of Meyer and Trauth decrease relative to the control group prediction significantly with distance. In particular, their bids are significantly lower in the distance ranges where they were unusually likely to bid—Meyer at 100 to 110 miles, and Trauth at 60 to 80 miles. As further evidence, consider bid level regressions, comparable with the control group regressions, for the three Cincinnati dairies. For both Meyer and Trauth, which submitted distant bids, bid levels are a significantly decreasing function of distance. In contrast, the distance bids are significantly higher for the control group.

To focus on whether firms behave in a parallel fashion, we tested for statistical independence in the probability of bidding for the defendant firms using a standard pairwise procedure. Under the null hypothesis of independent action based on public information and the specifications of our bid submission model, knowledge of whether one particular firm bids should not help predict whether another firm also bids. Under an alternative hypothesis of either complementary bidding or territorial allocation, the

---

\(^1\) Hendricks and Porter (1989) describe several collusive mechanisms, and why the detection of collusion is necessarily case specific. Klemperer (2008) explains why auction market competition policy should differ from other markets.

\(^2\) By "actual bid," we are referring to the summary bid measure, rather than any element of the vector of bids that the firm may have submitted.
submission decisions are interrelated, and knowing how one cartel member behaves helps predict what the other does. In the case of complementary bidding, if one cartel member bids, then other ring members also bid. In this case the unexplained portion of the competitive bidding equation is positively correlated across cartel firms. In the case of territorial allocation, if a particular cartel member bids, then other cartel members will tend not to bid. Then the unexplained portion of the competitive bidding equation is negatively correlated across cartel members. The test for independence, or zero correlation, that we used has power against both alternatives. The results indicate that the unexplained portion of the Coors submission decision was positively correlated with the unexplained portion of the Meyer decision, and similarly for Coors and Trauth, and for Meyer and Trauth. All three correlations are significant at any standard level.

We perform a similar analysis for the level of the submitted bids. Under the null hypothesis of independent action based on public information and the maintained specification of the bid level equations, knowledge of what one particular firm bid does not help predict what another firm will bid. Under an alternative hypothesis of complementary bidding, knowing that one cartel member bid above the predicted level helps predict whether other cartel firms will bid above that level. If one cartel member bids high, then other ring members are also likely to bid high. The pairwise correlation coefficients are positive and significant for all three pairs.

Our results support the testimony by representatives of Meyer and Coors. The behavior of Coors, Meyer, and Trauth is consistent with a complementary bidding scheme in the area close to their plants, since more bids than expected are submitted at distances of less than thirty miles. Further, these bids tend to be relatively high. The correlation results for these three firms are also consistent with a complementary bidding scheme. There are statistically significant correlations in bid submissions by these firms, suggesting that if one of these firms bids, then the others also tend to bid (to a greater extent than their proximity, size, and other factors would predict). In addition, when these firms bid on the same districts in the same years, their bids tend to move together to a greater extent than their proximity, size, and other factors would predict.

It is difficult to craft a competitive story where bid levels decrease with distance, as they do for these firms. The behavior of the Cincinnati dairies is suspicious, even without a comparison with the behavior of nondefendant firms. We believe that the collective behavior of these three firms is best characterized as collusive.

The Defendants' Response

Not surprisingly, the defendant dairies disagreed with our interpretation of the evidence. They criticized our econometric approach on several grounds. For example, they noted that the list of explanatory variables in
our econometric model of bidding behavior was incomplete, in that they
could point to a number of other factors that might also affect the decisions
of whether to bid, and at what level. However, they did not provide a theo­
retical argument or any econometric evidence to indicate that the effect of
omitting these variables would necessarily bias our tests in any particular
direction. Instead, the force of their argument was that the model did not fit
as well as one might like, and our conclusions therefore were suspect.

Second, they argued that the control group was not a cohesive group,
because one could reject the hypothesis that they all bid according to
exactly the same bidding model, and therefore a comparison of the defen­
dants to the average behavior of the nondefendants is not an interesting
exercise. However, Porter and Zona (1999) report that the bidding behavior
of all but three of the fifteen nondefendant firms (those for which there are
sufficient data to conduct the test) is not significantly different than the
average behavior of the remainder. This contrasts with the defendants, all of
whom have significantly different behavior.

Finally, the defendant dairies argued that econometric evidence
alone cannot distinguish between an overt conspiracy and tacit collusion,
whereby the dairies charged high local prices without ever communicating
with each other.

EFFECT OF COLLUSION ON PRICES PAID

On the assumption that a conspiracy involving all the southwestern Ohio
defendants was in force throughout the 1980s, what were the likely dam­
ages? Our methodology for estimating damages to the plaintiff school dis­
tricts involves determining the percent markup in price attributable to
collusion in various auctions.

Using regression techniques, Porter and Zona (1999) estimated the
effect of variables that determine costs for the most efficient provider and
other variables on the standardized price of the winning bid. The sample
covers about four hundred districts, including those outside of the south­
western region. Annual dummy variables control for changes in the raw
milk price, changes in uncertainties in the raw milk price over time, and
changes in the costs of packaging year to year. We also control for district
enrollment, the number of deliveries, and other characteristics of the school
district. We include two location variables, to account for the effects of
(1) the distance from the closest plant to the school district, as we expect
price to be increasing with distance to the closest plant as shipping costs
increase, and (2) the distance between the district and the second-closest
plant. When the latter distance is large, all else equal, we expect the price
paid by the school district to be higher since the closest firm can charge a
higher price.
We also control for the effects of differing levels of competition in each of the markets. We measured competition in these markets using the number of equivalent firms defined by the inverse of the Herfindahl index. We expect higher prices in more concentrated markets where fewer firms compete with one another.

If there is collusion in these markets, prices are on average above the competitive level. Therefore, we include in each regression an index of collusion based on the number of conspirators in this case who are within seventy-five miles of the school district in question. We also include an interaction between the collusion index and the number of equivalent firms, since the effect of a restriction of competition will depend on the initial level of competition. We interact these indices of collusion with the annual dummy variables so as to measure, to the extent they exist, differences in the degree of collusion from year to year.

In our regression we predicted the winning bid using 2 percent chocolate milk as the base product. In general, the coefficients are of the anticipated signs and statistically significant. The number of firms in the market, as measured by the variables described above, indicates the significant effect of concentration on the price paid by the school districts.

The predicted increase in price caused by collusion in these auctions is measured as the difference between the predicted value when indices of collusion are included and the predicted value that is obtained when all indices of collusion are set to zero. The difference is the predicted percentage change in price resulting from collusion. The estimated effect of collusion is small in 1983-1984 and after 1989, consistent with the market participants' testimony that the cartel agreement broke down in those years.

The average effect of collusion on price is an increase of about 6.5 percent. This is consistent with our estimate of the average effect of distance on school milk bids. If, for example, two nearby firms conspired to serve a district adjacent to their plants and faced competition only from firms located at least fifty miles from that school district, then prices could be about one-half penny (or about 5 percent) higher than they would be otherwise. Districts that are farther from potential competitors face higher markups, and districts closer to the plants of competitive firms pay lower markups.

The Herfindahl index is based on the fraction of school milk processing plants within seventy-five miles of the district controlled by each processor. If there are ten owners of ten processing plants within seventy-five miles, then the Herfindahl is one-tenth, and the number of equivalent firms is ten. If there are ten dissimilar firms, the Herfindahl index will exceed one-tenth, and the number of equivalent firms will be less than ten.

The index of collusion is the difference between the inverse of the Herfindahl index assuming competition and the inverse Herfindahl assuming collusion.
SUMMARY AND AFTERMATH

Our analysis exploits specific features of school milk markets, and we would not necessarily advocate using similar methods to study other auction markets. For example, an unusual feature of Ohio school milk auctions is that few bids are submitted, even relative to school milk auctions in other states. An important component of strategic decisions in this market is whether to submit a bid.

We also focus on the role of distance in bidding decisions. Processed milk is relatively expensive to ship, and competition is localized. We emphasize the fact that two of the three Cincinnati dairies tend to submit relatively low bids for distant school district contracts, and yet submit higher bids close to their plant. We document these patterns with a bidding model that ignores characteristics of potential rivals, except to the extent that we control for district specific fixed effects, in large part because we cannot be sure whether other firms were indeed rivals. In the case of the Cincinnati dairies, this omission should bias the results in their favor. Had they competed aggressively, bids for nearby districts should have been relatively low, as there were three local potential suppliers. Yet, prices in Cincinnati were relatively high, especially in comparison with the dairies’ distant bids. This price pattern is consistent with local monopoly power, but local monopoly power is consistent only with collusion.

The detection of collusion in auction or other markets is necessarily case specific, since the details of a collusive scheme depend on a number of factors. The study of auction markets can nevertheless inform the study of collusion in other markets. Data are often available on the bids of all participants in an auction, and the auction rules define the set of strategic considerations. Under these circumstances, a detailed study of strategic behavior is feasible, and it may be possible to isolate behavior that is inconsistent with competition.

This case study describes some features of a conspiracy among a group of neighboring suppliers, when the conspirators are protected locally from competition by the transportation costs borne by distant suppliers, and when the conspirators compete on a relatively equal footing with other firms outside their local market. In this case, lower prices at more distant locations are consistent with a local conspiracy and inconsistent with competitive behavior.

The case brought by the state of Ohio against the Cincinnati dairies was settled in 1996 before it could go to trial. Earlier, the federal government lost a criminal conspiracy case in 1995 that centered on the testimony of the Meyer and Coors representatives. In the federal trial, the defense described this testimony as unreliable, on the grounds that it contradicted previous statements of innocence before a grand jury. The federal case did not rely on statistical or economic evidence. The state preferred to settle rather than to go to trial with the same direct evidence, which would be
subject to similar questioning. The economic evidence was intended to corroborate this testimony and to assess the amount of damages. It is very difficult to distinguish overt conspiracy from tacitly collusive behavior with economic evidence, which is typically not regarded as sufficient to establish liability.

In 1993, the U.S. Department of Justice (DOJ) clarified its Corporate Leniency Policy, under which the first cartel member to confess participation is awarded complete amnesty from prosecution (unless it was already the subject of an investigation, in which case amnesty is limited). All other participants are subject to harsh penalties. DOJ officials believe that this policy is now their most effective weapon in combating cartels (e.g., Hammond, 2001), one that induces a "race to the courthouse." Recent government cases have relied heavily on the testimony of participants in the conspiracies. The 1993 leniency policy did not shield informants from civil penalties, most notably from treble damages, but the 2004 Criminal Penalty Enhancement and Reform Act limits informants' penalties to actual damages (plus attorney fees) and increases the potential liability of other conspirators.

Rey (2003) and Spagnolo (2008) describe how leniency programs are implemented in the United States and Europe, and they discuss how leniency policies should be designed to fight collusion effectively. An effective policy would hasten the demise of existing conspiracies, as the costs of defecting from the agreement decrease, and the gains associated with continuing are limited. An effective policy would also prevent some new cartels from forming, if potential conspirators fear the potential legal consequences.

REFERENCES


PART III
Vertical and Related Market Issues
The Economic and Legal Context

Vertical and complementary market relationships between firms continue to be areas of change for both microeconomics and antitrust policy. The issues involve firms that either are in a customer-supplier relationship with each other (vertical) or that sell complements; as will be argued below, the latter type of relationship is often analytically similar to the former, and we will treat them similarly.

There is a history prior to the 1970s, for both law and economics, of muddy thinking and muddled analysis in this area (White 1989). At least part of the problem seems to have been difficulties in conceptualizing the relationships and in recognizing their fluidity; even terminology (e.g., "vertical restraints," "foreclosure," "refusal to deal") may have played a role as well.

We will first address the economics of these relationships, and then address the law.

ECONOMICS

Some Fundamentals

A vertical relationship describes the business relationship between a supplier and a customer. Thus, a wheat farmer and a flour miller, the flour miller and a bread baking factory, the baking factory and a supermarket, and so on, are each in a vertical relationship with the other. Often these relationships are described as "upstream" (the supplier) and "downstream" (the customer) relationships. Another way of describing them is that each party in principle is always facing a "make or buy" decision with respect to the goods or services of the other.

Some References:

This last point highlights the possibility of *vertical integration*: Vertically related activities that are (or could be or formerly were) located in separate (customer-supplier) businesses could be combined and integrated "under one roof" in a single enterprise. Thus, the farmer might decide to mill his wheat himself and then sell the flour to the baker, or the miller might decide to grow his own wheat; equivalently, the baker might decide to open its own retail outlet, or the supermarket might start baking its own bread.

In turn, these illustrations highlight the fluidity of vertical arrangements. In principle, almost any vertical relationship between firms could be vertically integrated into a *single* entity, and almost any vertically integrated entity could be broken into two separate, vertically related firms. There is thus no "natural" extent of vertical integration for a firm; which of its inputs it "makes" or "buys" is always a situationally specific issue, driven by factors that we will discuss below.

Further, even the question of who is the "upstream" party and who is the "downstream" party is situationally specific. For example, manufacturers typically sell their goods to distributors or retailers, who in turn sell to the public, but sometimes manufacturers "buy" distributional services (e.g., through the payment of fees and commissions) while retaining title to the goods until they are sold to the public. Who is "upstream" and who is "downstream"?

At the opposite pole from vertical integration is the spot market. Here buyers and sellers each appear on an *ad hoc* basis, determined by their respective needs to participate in the market and engaging in one-at-a-time transactions.

In between the two extremes is a wide range of vertical arrangements between buyers and sellers that are agreed to by both parties and that limit the behavior of one or both; these are often embodied in legally enforceable contracts and are frequently termed *vertical restraints*. By thereby giving one party some control over the other party's actions, these arrangements are, in essence, a form of partial vertical integration. These arrangements include:

- long-term contracts;
- franchising;
- licensing;
- tying;
- bundling;
- exclusive dealing;
- requirements contracts;
- full-line forcing;
- refusals to deal;
- territorial restraints; and
- resale price maintenance.
Just as the extent of vertical integration is situationally specific, so is the presence or absence of vertical restraints.

Finally, firms are often in a relationship that appears to be neither horizontal (i.e., competitors) nor vertical (customer-supplier). For example, both the flour miller and the yeast manufacturer sell inputs to the bread baker. The relationship between the flour miller and the yeast manufacturer is neither horizontal nor vertical; they are selling complements. However, we could instead envisage an alternative structural relationship, whereby the flour miller sells his flour to the yeast manufacturer, who then sells a flour-yeast mix to the baker; for many analytical purposes (e.g., evaluating mergers), the two alternative arrangements yield similar outcomes. Accordingly, much of the discussion in this part applies equally strongly to vertical arrangements and to relationships between sellers of complementary goods.

**Why Vertical Integration?**

As the previous section indicated, vertical relationships embrace a broad spectrum of possibilities, from the vertical integration of activities within one firm to spot market transactions between firms, with various intermediate possibilities. Let us begin by considering the contrast between vertical integration and spot market transactions. Though there would seem to be advantages to the latter arrangement, reliance on the spot market may entail considerable informational and transactions costs (Coase 1937). Presumably, profit-seeking firms will search for the arrangements that offer them the most profitable opportunities. Whether this will involve "making" (i.e., vertically integrating) or instead "buying" will thus be driven in an immediate sense by perceived profit opportunities. But, in a larger structural sense, whether activities are integrated or in separate enterprises will be related to a host of potential factors, including:

- the technologies of the production processes;
- business strategies;
- managerial capabilities;
- customer preferences;
- legal requirements or restrictions;
- tax considerations;
- product-related informational issues;
- the presence (or absence) of market power; and
- the pursuit or protection of market power.

Vertical integration is often a means of taking advantage of technological complementarities, reducing transactions costs, gaining greater control over production processes (e.g., so as to preclude opportunistic behavior by
suppliers), overcoming informational deficiencies, and/or internalizing externalities. The problem of “successive monopolies” or “double margin-
alization” is a good example of this last phenomenon (Tirole 1988, ch. 4).

If an upstream monopolist sells to a downstream firm that is also a monop­
olist, the former firm will include its profit margin in its price; the latter firm
thus buys that input at a price that is in excess of marginal cost, which in
turn will cause the downstream firm to set a price that is excessively high,
and both firms’ sales volumes and profit levels suffer. The fundamental
problem is that of an externality (each firm ignores the effects of the other’s
market power), and vertical integration between the two firms internalizes
that externality and allows the combined firm to sell at a lower price that
also yields higher profits.’

But vertical integration may also be a means by which a firm creates or
enhances market power, by raising entry barriers and/or raising rivals’ costs.
A dominant upstream firm may be able to disadvantage its upstream rivals
by integrating downstream (e.g., through merger) and reducing its upstream
rivals’ access to distribution; or, equivalently, a dominant downstream firm
may integrate upstream and thereby disadvantage its downstream rivals by
reducing their access to suppliers.’ Also, vertical integration downstream
may be a means by which a firm can more effectively practice price dis­

The double-marginalization problem also provides a good example of the parallel between verti­
cally related firms and complementary sellers. If two complementary sellers are monopolists and
each ignores the other’s monopoly power, then the downstream price will be too high, both sellers’
profits will suffer, and a merger between the two would internalize the externality. This point is
raised in the discussion of the GE-Honeywell proposed merger in Case 14 by Barry Nalebuff in this
part.

Note that these scenarios imply that an input (including distribution) is essential and good substi­
tutes are not easily available.

The social welfare consequences of price discrimination, however, are ambiguous; see, for exam­

Since a profit-maximizing monopolist maintains its price at a level where (by assumption) a higher
level would be unprofitable, a price increase test (such as the SSNIP) for determining whether a
market can be monopolized should always find that even the true monopolist cannot profitably raise
its price from the current level (as is true also for a competitive firm). Thus, this test would mistak­
enly find that even the true monopolist appears to be just another competitor.
Why Vertical Restraints?

As was argued above, vertical restraints can be conceptualized as agreements that allow one or both parties to restrict the actions of the other and that thus can be considered partial vertical integration. The same forces—technology, managerial capabilities, and so on—that influence the presence or absence of vertical integration tend also to influence the presence or absence of these arrangements. Often a manufacturer will have strong ideas about the way in which its product should be distributed and promoted at the retail level, but does not have the managerial capabilities to undertake the retailing itself. Vertical restraints are a means by which these ideas can be implemented without the necessity of the less-efficient full vertical integration. However, though they often promote efficiency, vertical restraints may also be a means for attaining or enhancing market power.

The pluses and the minuses for various vertical restraints include the following:

• Resale price maintenance (RPM) involves the specification of the retail price by an upstream entity (e.g., the manufacturer). RPM that sets a minimum price can be a means by which a manufacturer can induce the provision of point-of-sale services, especially information, by its retailers, each of whom might otherwise be tempted to free-ride on the services of others; but RPM may also be a cover for a horizontal conspiracy among retailers or among manufacturers (Telser 1960).

• Territorial restrictions involve a manufacturer’s restricting the territories in which its retailers can locate or the locations of customers to whom they can sell. The restrictions can allow a manufacturer to avoid duplicate sales efforts and free-riding (e.g., on the advertising of others) among its geographically dispersed retailers; but the restrictions may also permit the manufacturer more effectively to practice price discrimination or to enforce a retailer-inspired conspiracy to allocate markets and restrict competition (White 1981).

• Long-term contracts between a manufacturer and a distributor provide both parties with greater levels of certainty, which can be valuable for

---

6 Often the efficiency rationale for these practices can be best comprehended by imagining how the relevant behavior would appear if the two firms that are parties to the vertical restraint were vertically integrated.

7 In the examples below we will use “manufacturer” to represent generally an upstream entity (supplier) and “retailer” to represent the downstream entity (customer).

8 RPM that sets a ceiling on prices may be a way to restrict the market power that is created for a retailer when a manufacturer also grants an exclusive territory to the retailer. See the discussion of the State Oil litigation in Case 17 by Gustavo Bamberger in this part.

9 Similar restrictions may apply to the category of customer to whom the distributor can sell, with similar rationales.
planning and investment (and help avoid “hold-up” or other opportunist­
tic behaviors by one side or the other) (Williamson 1983); but such con­
tracts may also restrict the access of rival manufacturers to distribution
services, thus raising their costs of distribution, or raise barriers to
entry (if distribution services are an important bottleneck for that line of
products).

- Exclusive dealing involves the manufacturer’s insisting that the retailer
sell only that manufacturer’s line of goods. The manufacturer may
thereby achieve the focused effort of the retailer and prevent free-riding
by other manufacturers; but the practice may also raise rivals’ costs or
barriers to entry, similar to the effects of long-term contracts.

- Refusals-to-deal involve the manufacturer’s choosing to distribute its
product through some retailers and not others. The practice permits the
manufacturer to enforce other types of restraints (by refusing to sell to a
retailer that does not abide by the restraint) and to avoid selling through
retailers who the manufacturer believes do not enhance the goodwill of
the product. But refusals-to-deal may similarly help the manufacturer
effect anticompetitive outcomes from other restraints or may simply
make some retailers’ existences more difficult (which may adversely
affect other manufacturers).

- Tying and bundling can be a means of ensuring quality and improved
overall performance of a product or service; but these practices can also
be used to price discriminate, to raise prices for the combined package,
or to raise rivals’ costs and barriers to entry.

- Franchising and other licensing arrangements can allow the originators
of an idea to extend its use and maintain quality control over its use, with­
out their engaging in the direct production (in which they may have little
relevant expertise) themselves; but these arrangements may also be used
to effect ties or other vertical restraints.

Further, as was true for vertical integration, the presence or absence of
actual or potential market power is crucial for judgments about these prac­
tices; and the issue of market definition is just as important and just as
unresolved.

"Equivalently, a retailer may insist that manufacturers sell only to it and not to other retailers. This
version is raised in the Toys "R" Us dispute that is discussed in Case 16 by F. M. Scherer in this part.

"See the discussion of the Independent Ink case in Case 13 by Barry Nalebuff in this part.

"See the discussion of the Kodak case in Case 21 by Jeffrey MacKie-Mason and John Metzer in
Part IV.

"See the discussion of the Microsoft case in Case 20 by Daniel Rubinfeld in Part IV.

Page 357
ANTITRUST LAW

A broad range of the antitrust laws' provisions has been used to attack vertical and complementary market arrangements:

- The Sherman Act's Section 1 prohibition on any "contract, combination ... or conspiracy in restraint of trade ..." has been used to challenge vertical restraints.
- The Sherman Act's Section 2 condemnation of "monopolization" has been used to challenge vertical restraints and vertical integration itself.
- The Clayton Act's Section 3 specifically forbids contracts imposing a restraint whereby the customer "shall not use or deal in the goods, supplies, or other commodities of the lessor or seller" where the effect "may be substantially to lessen competition or tend to create a monopoly."
- The Clayton Act's Section 7 (which was discussed in Part I) has been used to challenge vertical mergers.

For the discussion that follows it is useful to classify the antitrust treatment of vertical arrangements into four areas: vertical integration; restraints involving prices; territorial restraints; and other nonprice restraints.

Vertical Integration

Section 2 of the Sherman Act has been used, albeit sparingly, to attack vertical integration that appeared to promote or enhance market power. The 1911 Standard Oil decision resulted in vertical dissolution (e.g., oil pipelines being separated from refineries) as well as horizontal dismemberment (refineries in different geographic regions were separated into independent companies). The 1948 Paramount case resulted in the separation of movie exhibition (theaters) from distribution and production. The Justice Department's AT&T case and 1982 consent decree, which separated long-distance and equipment from local telephone service, was grounded on the DOJ's theory that AT&T's vertical integration encompassing local service, long-distance service, and equipment manufacturing, when buttressed by imperfect local and national economic regulation, created competitive distortions (Noll and Owen 1994).

Section 7 of the Clayton Act has been used to prevent vertical integration through merger. Even before Section 7 was strengthened in 1950, the DOJ was able successfully to challenge General Motors' substantial stock ownership position in its major supplier of paints and finishes.

U.S. v. Paramount Pictures et al., 334 U.S. 131 (1948).
In the heyday of merger prohibition victories of the 1960s, vertical mergers (as well as horizontal mergers) were regularly challenged and stopped. Fewer vertical mergers received scrutiny in the 1980s and 1990s, as the enforcement agencies came to understand the more favorable economic assessment of these mergers' effects. Nevertheless, some mergers did require modifications before they were approved, because of the agencies' concerns about the creation of horizontal market power, through the strategies of raising rivals' costs and/or raising barriers to entry (White 1985; Nelson and Stoner 1999; Besen et al. 1999). And in 2001 the Department of Justice (DOJ) showed its continuing attention to vertical issues by challenging the acquisition by Premdor Inc. (a manufacturer of finished residential molded doors) of the Masonite Corp. (a manufacturer of "doorskins," the essential molded input for these doors) and requiring the divestiture of some facilities before approving the acquisition (Katz 2002).

**Resale Price Maintenance (RPM)**

In the 1911 *Dr. Miles Medical Co.* case the Supreme Court declared that RPM was a per se violation of Section 1 of the Sherman Act, thus placing RPM in the same condemned position as horizontal price fixing. The Court repeatedly upheld that position in subsequent cases, although in the 1980s it raised the evidentiary standards for finding a defendant guilty of RPM.  

An important new development in the 1997 *State Oil* case involved maximum RPM. In 1968 the Supreme Court had declared that maximum RPM was a per se violation of Section 1, in the *Dr. Miles Medical* tradition. But, as discussed in Case 13 by Gustavo Bamberger in this part, the Court decided in *State Oil* that maximum RPM should be judged under a rule of reason. And, finally, in *Leegin* in 2007, the Supreme Court decided that minimum RPM, too, should be judged by the courts under a rule of reason.

"U.S. v. E.I. Du Pont de Nemours and Co. et al., 353 U.S. 586 (1957)."
"See, for example, Brown Shoe Co. v. U.S., 370 U.S. 294 (1962); Reynolds Metals Co. v. F.T.C., 389 F.2d 223 (1967); and Ford Motor Co. v. U.S., 405 U.S. 362 (1972)."
"Dr. Miles Medical Co. v. John D. Park and Sons Co., 220 U.S. 373 (1911)."
"If some neighboring retailers complain to a manufacturer that retailer X is shirking on service and is cutting prices, and the manufacturer subsequently terminates that retailer, does that constitute evidence of a RPM scheme? In Monsanto Co. v. Spray-Rite Service Corp., 465 U.S. 752 (1984), and Business Electronics Corp. v. Sharp Electronics Corp., 485 U.S. 717 (1988), the Supreme Court declared that more than just this fact pattern would be necessary for conviction. See Warren-Boulton (1999)."
"State Oil v. Khan, 522 U.S. 3 (1997)."
"It is noteworthy that the DOJ and the Federal Trade Commission (FTC) entered a joint amicus brief in the State Oil case, urging this outcome."
"Leegin Creative Leather Products, Inc. v. PSKS, Inc., 127 S.Ct. 2705 (2007). In this case, too, the DOJ and the FTC joined in urging the Supreme Court to adopt the rule of reason approach."
Territorial Restraints

The Supreme Court first addressed territorial (and customer classification) restraints in 1963 and declared that it felt that not enough was known about them to condemn them as per se violations of Section 1 of the Sherman Act. 25 Four years later, however, the Court reversed itself and declared such practices as per se offenses. 26 In 1977, however, the Court reversed itself again and declared in GTE-Sylvania 27 that these practices should be judged under the rule of reason. 28 The Court’s position on this issue has remained unchanged since then.

Other Nonprice Restraints

Until the mid-1970s the Supreme Court and the enforcement agencies were generally hostile toward tying, exclusive dealing, and other nonprice vertical restraints. In 1949, for example, the Court declared that “tying agreements serve hardly any purpose beyond the suppression of competition,” 29 and in 1958 it declared tying to be a per se offense. 30 Long-term contracts, requirements contracts, and other restraints were routinely found to be illegal by the courts, largely based on their “restrictive” nature. Since the 1970s, however, the approaches of the courts and the enforcement agencies have become more sophisticated. The Supreme Court has added successive exceptions and evidentiary burdens to plaintiffs’ efforts to win a tying case (but the Court still maintains that tying is a per se offense), 31 and the other nonprice restraints are judged under the rule of reason. The Supreme Court clearly stepped back from this trend in its Kodak decision. 32 But as the discussion in Case 21 by Jeffrey MacKie-Mason and John Metzler in Part IV indicates, the decision involved Kodak’s motion for summary judgment, 33 and not the explicit merits of tying, and the Court considered sophisticated economic arguments from both sides before reaching its decision. And, in Independent Ink, 34 discussed by Barry Nalebuff in Case 13

28 A discussion of this case, and the cases that led up to it, can be found in Preston (1994).
31 In its unanimous dismissal of a tying case in 1984, the Court barely failed (by a 5-4 vote) to eliminate tying’s per se treatment and adopt a rule-of-reason approach. See Jefferson Parish Hospital District No. 2 v. Edwin G. Hyde, 466 U.S. 2 (1984); see Lynk (1999) for a discussion of this case.
33 In essence, Kodak was arguing that even if all of the facts alleged by the plaintiffs were true, the plaintiffs as a matter of law could not win their case. The Supreme Court decided that Kodak had not sufficiently supported that claim.
in this part, the Supreme Court decided that the presence of a patent on a tying product does not automatically create the presumption that the seller of that product has market power (which would invoke the per se illegality rule for tying).

The enforcement agencies have moved away from a formalistic focus on "restrictions" and "foreclosure" and instead have focused their attention on those practices that seem likely to raise rivals' costs, raise barriers to entry, or otherwise exclude rivals. For example:

- The DOJ’s first investigation and consent decree with Microsoft in 1995 (Gilbert 1999) focused on Microsoft's practices that raised barriers to entry in PC operating systems; the DOJ’s subsequent challenge to Microsoft, which is discussed in Case 20 by Daniel Rubinfeld in Part IV, focused on Microsoft's efforts to protect its market power in PC operating systems by raising rivals' costs.
- The FTC’s dispute with Intel, discussed by Shapiro (2004), featured the FTC’s concerns that Intel's refusal to share intellectual property could raise barriers to entry.
- The FTC’s prosecution of Toys “R” Us, discussed in Case 16 by F. M. Scherer in this part, highlighted the agency's concern that Toys "R" Us’s actions (insisting on preferential treatment by toy manufacturers) raised rivals’ costs.
- The DOJ’s suit against Visa and MasterCard, discussed in Case 19 by Robert Pindyck in Part IV, involved the agency’s concerns that the two card associations' prohibitions on their member banks' issuing any other credit cards except Visa and MasterCard raised rivals' costs and barriers to entry.
- The DOJ’s prosecution of Dentsply International, Inc., a manufacturer of artificial teeth, discussed in Case 14 by Michael Katz in this part, claimed that its policy of refusing to sell its teeth to dealers that carried other manufacturers' teeth raised rivals' costs.

In sum, vertical relationships are unlikely any time soon to be a settled area of economics or of antitrust law. As we argued above, vertical arrangements between firms are always situationally specific. And if changes in technology or consumer preferences or other environmental

"One concept that the Antitrust Division has been promoting in its filings in vertical restraint cases is the "no economic sense test." The basic idea is to try to identify whether the defendant's actions would have made economic sense in the absence of the exclusionary effects that are posited. For an exposition of this position, see Ordover and Willig (1981, 1999) and Werden (2006); for a critique, see Salop (2006).

"As MacKie-Mason and Metzler point out, the Kodak case alone has been followed by numerous circuit court opinions on similar issues, with deeply divided results."
conditions lead a firm to decide to change its vertical arrangements, there may well be customers or suppliers or rivals who are disadvantaged by the change and feel aggrieved. Lawsuits may well follow, and there will almost always be credible efficiency arguments that can be offered to counterbalance any anticompetitive claims. Creative economic theorizing about the pluses and minuses of vertical arrangements will surely continue. And a set of economically sound principles for delineating markets in such cases, so as to make sensible judgments about the presence or absence of market power, is absent.

Vertical arrangements will remain a rich stew indeed.

REFERENCES


“It is noteworthy that the Antitrust Modernization Commission (2007, ch. I.C), in its discussion of “exclusionary” conduct, focused largely on bundled discounts and refusals to deal and did not make any recommendations with respect to tying or RPM.”

Page 362


INTRODUCTION

The case of Illinois Tool Works, Inc. and Trident, Inc. v. Independent Ink, Inc., 547 U.S. 28 (2006), generated a good deal of controversy. The Supreme Court used this case to overturn a longstanding per se rule against tied sales by a firm with a patent. Henceforth, market power will have to be demonstrated, whether or not the firm has a patent. Upon demonstration of market power, a tied sales contract will still be a per se antitrust violation.

At one level, the Court’s decision reflected an obvious truth: A patent alone does not demonstrate market power. At a deeper level, this case concerned why firms engage in metering and tying, and what is the antitrust issue with these practices. While the courts have often focused on the leveraging of market power, this case suggests a different concern: With complementary goods (such as a printer and ink), tying is used to engage in price discrimination via metering. The antitrust issue is that consumers will be harmed. This is because the firm, by metering, is able to be a more effective monopolist and thereby extract more of the consumer surplus.

According to Winston Churchill, “History is written by the victors.” If so, I should not be the one writing this chapter. Together with Ian Ayres and Lawrence Sullivan, I had the opportunity to present an amici curiae brief to the Supreme Court in this case (Nalebuff et al. 2006). In early 2006, the U.S. Supreme Court rejected our arguments by a vote of 8-0. Justice O’Connor had just stepped down from the bench, thus sparing us from a 9-0 result. This chapter was helped by feedback from Ian Ayres, Dan Bromberg, Barry Brucker, Walter Grimes, Alan Horowitz, Kathleen Sullivan, Lawrence Sullivan, and the editors John Kwoka and Lawrence White.
Thus, the more difficult questions raised by this case are the following:

- Is a metering contract by itself evidence of market power? Is price discrimination compatible with competition?
- Does price discrimination via metering harm consumers and efficiency?
- What is the potential efficiency loss of prohibiting metering via tied sales (for firms with market power)?

BACKGROUND

Trident, a subsidiary of Illinois Tool Works, is a leading manufacturer of piezoelectric printheads. These components are used in high-speed industrial inkjet printers that print bar codes, graphics, and other text onto blank cartons as they move along a production line. Printer manufacturers obtain Trident's patented printheads under license agreements. The agreements provide for the ongoing exclusive purchase of ink from Trident, and the exclusivity agreement extends to the ultimate buyer. Independent Ink is a small manufacturer of ink whose product is compatible with Trident printheads. Independent Ink's equivalent product sells for between $125 and $189, a steep discount to Trident's charge of $325. This price discount attracted manufacturers to use Independent Ink's ink, in violation of their exclusivity agreement.

Trident responded by threatening to cut off printer suppliers and to sue the end-use customers. Then, in 1997, Trident sued Independent Ink for patent infringement in Illinois. The case was dismissed for lack of personal jurisdiction as Independent Ink had little commercial presence in that state.

In 1998, Independent Ink sought a declaration that its inks did not infringe Trident's patents and that Trident's patents were invalid. Trident countersued for patent infringement, and Independent Ink then amended its complaint to include claims of monopolization and illegal tying.

The Federal District Court for Central California dismissed Trident's patent infringement claim. At the same time, the district court granted summary judgment for Trident denying Independent Ink's antitrust claims. The basis for the summary judgment was that Independent Ink had not provided a traditional economic analysis demonstrating that Trident had market power. Independent Ink relied on a per se rule based on the presumption of market power in tying cases involving a patent. The district court rejected Independent Ink's submission that Trident "necessarily [has] market power for inkjet printhead components in the United States".

in the market for the tying product as a matter of law solely by virtue of the patent on their printhead system, thereby rendering [the] tying arrangements *per se* violations of the antitrust laws (at 1159).

The Court of Appeals for the Federal Circuit agreed with the district court that a patent does not necessarily confer market power but also recognized that this conclusion was contrary to the prior six decades of Supreme Court precedents, from *International Salt* (1947) and *Loews* (1962) to *Jefferson Parish* (1984). While it was critical of those precedents, the court of appeals argued that it was not the role of the district court to overturn the Supreme Court. That was the job of the Supreme Court. This reversal of the summary judgment motion was appealed to the Supreme Court, which decided to hear the case.

Ultimately, the Court decided that it would overturn its own precedent. At the same time, it recognized that Independent Ink had reasonably relied on these precedents when making its case for a *per se* violation. Thus, Independent Ink was given an opportunity to demonstrate that Trident did indeed have market power and thus violated § 1 of the Sherman Act through its tying contracts.

The general interest in this case went well beyond the conflict between Trident and Independent Ink. The larger question concerned the ability of a firm to exercise and extend market power via tying. By tying we mean that customers were forced to accept a contract to buy some other good B along with the firm’s good A. Prior to this case, the law had two strands. According to the first, a tying contract was a *per se* antitrust violation for a firm with market power. According to the second strand, a tying contract was a *per se* antitrust violation for a firm with a patent. The law presumed that a patent conferred market power. Now there is just one strand: Market power must be established, patent or no patent. Going forward, even the first *per se* rule is under attack. Critics such as Evans (2006) and Wright (2006) argue that *Jefferson Parish* should be overturned and all tying contracts should be evaluated on a *rule-of-reason* basis.

We first review the arguments presented by the two sides. This leads us to analyze the motivations for a tying contract with an essential complement (e.g., a printer and ink). Tying creates an opportunity to meter usage and thereby increase profits through price discrimination. Although quality assurance, cost savings, and risk sharing might provide additional justifications, none explains why this type of contract is forced on the customer. This leads
to the conclusion that price discrimination is the true motivation. We then show that contrary to the Court's finding, this type of price discrimination should be viewed as evidence of market power. Such price discrimination will generally harm consumers and, contrary to the conventional wisdom, may also reduce overall efficiency. This suggests that tied contracts do create an antitrust concern and argues for the preservation of the per se rule against tying by firms with market power, at least where tying is used for metering.

ARGUMENTS

Trident, with the support of a brief filed by the U.S. solicitor general, argued that a patent, by itself, was not evidence of market power. On this point there was little controversy. Most patents are of no economic value. Since a patent does not ipso facto imply market power, this suggests that market power must be demonstrated in all tying cases.

Trident also argued that maintaining the per se rule in the case of patents would lead to a significant number of nuisance cases, imposing a burden on defendants. In his oral argument before the Supreme Court, the U.S. deputy solicitor general emphasized that the presumption that a patent creates market power in antitrust was based on a similar presumption in the patent code. But that presumption no longer exists.

Congress, in 1988 in the Patent Misuse Reform Act… held that there cannot be patent misuse in the absence of an actual showing of market power. So the rationale and underpinnings of Loew’s have been entirely repudiated...

Eliminating the presumption of market power in the case of antitrust would harmonize the patent law with the antitrust law.

Independent Ink countered that Congress had an opportunity to remove the presumption from the antitrust laws but did not do so. As Independent Ink explained during the oral argument:

And while congressional inaction might not always be a good guide to what Congress is thinking, here the Senate actually placed legislation in

This argument is questionable. Given the per se injunction, very few firms with patented products were willing to engage in a forced tie. Indeed, the Trident case was so unusual because the firm imposed the tied sale in such a clear and open fashion in spite of the existing law. A better objection would be that firms won’t be able to engage in tied contracts for legitimate purposes. We address this concern in the text.

The 1988 Patent Misuse Reform Act makes it clear that absent a demonstration of market power, a patent holder may condition the sale of the patented product on the purchase of a separate product. Note, however, that Trident's requirements tie for ink extended for the life of the printhead, not just the life of the patent. This specific type of requirements tie could be viewed as evidence of market power and a misuse of the patent.
the bill that was sent to the House to remove the presumption from the antitrust laws as well, and the House took it out and the Senate acquiesced.

Here the Supreme Court sided with Trident in regard to the interpretation of the new patent law (U.S.C. § 271(d)(5)). According to the decision:

[The amended code] makes it clear that Congress did not intend the mere existence of a patent to constitute the requisite ‘market power’. ... It would be absurd to assume that Congress intended to provide that the use of a patent that merited punishment as a felony would not constitute ‘misuse.’ Moreover, given the fact that the patent misuse doctrine provided the basis for the market power presumption, it would be anomalous to reserve the presumption in antitrust after Congress eliminated its foundation."

Even if the mere existence of a patent did not imply market power, Independent Ink argued that only the cases in which the patent did confer market power would lead to litigation. Thus, there should be a rebuttable presumption of market power. This argument was developed in an amicus brief (Scherer 2006). While most patents are worthless, a small number confer great market power. In the case of the valueless patents, the patent holder has no ability to leverage the patent, and the market will not see tying contracts. The only tying cases that will come to court are those in which the patent was one of the cases in which market power was present. "Results from a study of German patents show that patents in litigation are between 11.2 and 42.6 times more valuable than other patents, all else equal" (Scherer 2006, p. 2). Defendants would be given the opportunity to prove that theirs was not a case in which the patent conferred market power. The presumption would be against them, but they could still escape a violation by demonstrating a lack of market power.

This author's amicus brief offered a complementary perspective. The Court did not have to reach a conclusion about all patent cases, only the one in Independent Ink. Here the tying contract was of a very specific nature; it was for an essential complementary product (ink) that must be used with the patented product (printhead) in order to get the patented product to provide the output for which it was purchased. Thus, the relevant question is whether a patent-protected firm that engages in a forced tie to complementary aftermarket products has market power.

In these situations, the imposition of a requirements tie is strongly indicative that the patent holder does have significant market power, thus justifying a rebuttable presumption that favors the challenger to the tie. The basic argument is that a requirements tie for an essential complementary good to a patented product should be presumed to reflect market power because its primary purpose is price discrimination via metering. Such price discrimination is possible only when a firm has market power.

The Court addressed the question of whether a patent plus price discrimination should create the (rebuttable) presumption of market power. This position was debated during the oral argument and ultimately rejected in the decision. Following Baumol and Stanson (2003) and Levine (2002), the Court concluded that price discrimination is possible in competitive markets and thus there was no reason to presume that market power was a prerequisite for metering.

The process of overturning the per se rule against tying with a patent began when the district court made the observation that a patent did not automatically confer market power. The next logical step in this chain would be to conclude that tying does not automatically hurt consumers and thus should be subject to a rule of reason rather than a per se rule. The Court took the deputy solicitor general down this path during the oral argument. The antitrust rationale behind the per se rule against tying is suspect. The historical argument is that tying is used to leverage monopoly into the tied market. But, as Independent Ink makes clear, the motivation for tying, at least for complementary products, is more naturally price discrimination.

Thus, the most important antitrust question in this case concerns the potential harm caused by price discrimination, specifically metering. If this practice is viewed as benign, then there is no reason to prohibit tying contracts such as those required by Trident, even if Trident had market power. (Trident's brief argued that most tying arrangements are economically beneficial.) Conversely, if this practice is seen as harmful to competition and to consumers, then maintaining the per se rule is appropriate. This was the question that focused attention on this case.

Several commentators felt that the Court took a step in the right direction but missed the bigger opportunity to correct the bigger anomaly in the law; see Evans (2006) and Wright (2006). This is the issue to which we now turn. We provide an antitrust rationale for maintaining the per se rule against tying, even when it is used for price discrimination.

WHAT IS METERING?

Metering is a practice under which a customer is charged a price based on his or her use of the product. This allows the seller to raise its profits by charging more to high-value customers while still making sales at lower prices to low-value customers. It has long been understood that metering can be a useful price discrimination tool; see Bowman (1957).

Metering can be done directly or indirectly through a tied sale. For example, Monsanto charged a per-acre technology use fee for its patented Roundup ready seeds. Summit charged doctors a per-use fee for its patented laser eye surgery device.
Under indirect metering, the fee is based on the use of a complementary product, as with printers and toner or ink. For example, lessees of an A.B. Dick mimeograph machine were required to purchase all of their ink from A.B. Dick. Similarly, IBM required its tabulating machine customers to purchase its punch cards.

There are three salient features of a requirements tie used for indirect metering. First, the two products are essential complements in that the original product is of no value without the tied product. A mimeograph machine without ink is of no value. Second, the value of the original product is related to its intensity of use as measured through consumption of the complement. A customer who made more mimeograph copies would have generally placed a higher value upon the mimeograph machine and required more ink. A customer who did more tabulating would have generally placed a higher value upon the tabulating machine and thus have had a greater need for punch cards. More frequent use requires more complementary products, thereby allowing the seller to charge more to and earn more from high-value customers.

The final ingredient of indirect metering is that the tied aftermarket product is sold at a price premium. The price premium is how the firm profits from additional usage. As a result, the customer would prefer to buy the tied product elsewhere but is prevented from doing so either by contract, technology constraints, or lack of alternatives.

The potential for metering to increase profits is illustrated by the following example: Suppose that there are two consumers, each interested in a printer. All prints are valued at $2/copy. Customer 1 wants to make ten copies, while customer 2 wants to make only six. Absent a requirements tie, copies can be made at a competitive price of $1/copy (equal to the cost of the ink). Customer 1 will make ten copies that he values at $2/copy. Thus, consumer 1 expects to get a surplus of $1/copy and is therefore willing to pay up to $10 for the printer. Consumer 2 will make six copies that he values at $2/copy. Thus, consumer 2 is willing to pay up to $6 for the printer.

Without an option to engage in a tied sale or metering or the possibility of directly discriminating by charging $10 for the machine to customer 1 and $6 to customer 2, the monopolist maximizes profits by charging $6 for the machine, sells two machines, and earns profits of $12 (assuming zero production cost). When a requirements tied sale is possible, however,
the monopolist can force the buyers to use its overpriced ink. The ink is marked up to $2/copy, so that the monopolist earns $1/copy. Consumers are just willing to pay $2/copy but only if the machine is free as they earn no surplus. In this case, the monopolist can give away the machine for free and still earn higher profits. Customer 1 will make ten copies, and customer 2 will make six, leading to total profits of $16.

Note that the low user, customer 2, pays the same amount in both cases. For that customer, the reduction in the machine price is equivalent compensation for the increased cost of the tied good. However, customer 1 is now paying $10 rather than $6. That customer receives no equivalent compensation and hence is worse off, though not so much worse off as to reject the purchase of the machine.

This example illustrates that a monopolist charging a single price in the market does not fully exercise its market power. The use of a requirements-tied sales contract allows a seller to exercise its market power more fully.

There has been some confusion in both economics and the law as to whether tying can ever lead to additional monopoly profits. In Grappone, Inc. v. Subaru of New England, Inc., the appeals court stated:

> If the seller does have, and has been fully exercising, market power, it also cannot force buyers to take a more expensive or less desirable Product B, unless it provides buyers equivalent compensation by lowering the price of Product A...”

To the extent that this statement suggests that a monopolist cannot obtain more profits via a tied sale, it misreads the Chicago School one-monopoly profit argument.” That argument assumes that all consumers purchase identical quantities of good B, and thus there is no opportunity to engage in price discrimination via metering (Director and Levi 1956).

In our example, the monopolist increased its monopoly profits via the metering arrangement. There is still only one monopoly profit, but the size of that monopoly profit is variable based on the firm’s ability to engage in price discrimination. The monopoly profit under a single price was $12, while the monopoly profit under price discrimination rises to $16. While the monopolist does lower the price of the printer, the end result is not equivalent compensation for all buyers. The high-volume users, who had the most surplus, end up worse off under the metering contract.

As a result, aggregate consumer welfare is reduced by the metering contract. This is to be expected. The reason why consumer surplus typically falls is that those consumers with the highest value and the highest surplus are the ones who end up paying the biggest increase in price. The ability to

---

858 F.2d 792, 795 (1st Cir. 1988).

See the discussion below and Bork (1984).
price discriminate does expand the market, but the new customers are the ones with the smallest surplus to begin with, and they keep little of it. The largest source of increased profits is the transfer from consumers to the producer.

To the extent that a firm is given a legal monopoly via a patent, that monopoly power is expanded when the firm is allowed to engage in tied sales for the purpose of metering. The expansion does not arise because of power in the complementary market. Rather, the increased power comes in the original market.” While profits rise, consumers lose.

It is worth emphasizing that if the market for printers were competitive, such metering contracts would not be sustainable. Consider the position of the incumbent who is giving away printers with the requirement that its ink be purchased at a price of $2/copy. A rival would enter with an offer to sell an equivalent printer at $5 and let the customer buy ink at the competitive price of $1/copy. This rival would steal away both customers. The rival is best able to attract the most profitable customers by allowing those customers to use the machine most efficiently. This leads to selling the complementary good at its cost.”

Is Metering Evidence of Market Power?

It would appear from our example that the ability to engage in price discrimination via metering is evidence of market power. When a firm is able to price discriminate, it follows that some of its customers are more profitable than others. In a perfectly competitive market, such results are not possible. A rival firm that chose to undercut the price to the most profitable customers would gain all of those customers. Thus, even if firms are earning zero profits in aggregate, a firm could make money by targeting the most profitable segment of the market.

In the case of Independent Ink, customers who were the largest volume users of the inkjet printers were the most profitable. Were a rival to have offered an equivalent printer at a higher price combined with ink sold at marginal cost, this would have been attractive to the large-volume customers. The fact that a firm could profitably enter with a single-price contract (but none did) is evidence that Trident was exercising market power.

"One could interpret the passage in Grappone to say that a one-price monopolist is not fully exercising its monopoly power. The ability to price discriminate via metering allows a monopolist to exercise its power more fully. The point that this misses is that the court dismisses the view that customers lose as a result of being forced to buy the overpriced tied good. Here we see that this is the likely result.

"Later we discuss other potential motivations for metering, such as quality control. These arguments do not explain why the tying contract is required rather than optional and why the complementary good is priced well above its cost.
Is there a difference between market power as used in economics and market power as used in antitrust?” With some degree of product differentiation, a firm is able to charge a price above marginal cost. Does that mean that the local barber has market power when he charges a price above marginal cost? Yes, but barely. The barber has only a small amount of power in a local market. This tiny amount of power is irrelevant for antitrust. The markup is small, and no one would define the street corner as the relevant market for haircuts. The degree of market power is evidenced by how much the price rises above marginal cost and the scope of the market. It is also evidenced by how the firm is able to maintain its position.

In United States v. E. I. du Pont de Nemours & Co., the Court defined market power as “the power to control prices or exclude competition.” The metering contract is evidence of both. Trident was not able to charge a high price for its ink because the ink was differentiated or superior in quality. Customers wanted to purchase ink from Independent and were prohibited from doing so by their contract. Indeed, some customers were buying ink from Independent in spite of their contract.” Metering is not possible if the firm does not have the power to exclude competition.

Consider the experience of Wizard International. This company pioneered a machine to automate mat cutting for picture framers. When the machine initially came out, stores could rent it at a monthly fee and pay a charge per corner. Buying the machine was not an option. Wizard wanted to capture as much surplus as possible from high-volume users. But that approach created an opportunity for a competitor to enter the market. It could target Wizard’s most profitable customers by offering a fixed-price contract. Indeed, as Wizard faced competition, it changed its contract terms so that users could purchase its machine at a fixed price and pay no per-corner fees.

Technological innovations, often protected by patents, can give the firm market power that allows it to extract more profits by engaging in price discrimination. Such price discrimination leaves the high-profit customers open to competitors. When competitors enter the market, price discrimination is no longer sustainable.

This conclusion is not without controversy. Levine (2002) provides examples that lead him to argue that price discrimination is compatible with competition.” Baumol and Swanson (2003) go one step further and claim

“*The perspective below is influenced by Krattenmaker, Lande, and Salop (1987).*

“351 U.S. 377 (1956).”

“Levine observes that the price of sirloin exceeds the price of flank steak, and yet one is not more costly to produce than the other as the two are produced together in the form of a cow. Thus, the price differential reflects consumer preferences, not production costs, and is an example of price discrimination in a competitive market. A more careful examination shows that market prices do reflect costs. A firm would like to sell more sirloin and less flank steak. The problem is that a cow
not only that price discrimination is possible, but also that it may be required in some competitive markets.

The case of airlines is illustrative. The airline business would seem to be competitive based on the number of overlapping flights and the low profit levels of the industry. And yet price discrimination is rampant. Advance purchase restrictions and Saturday-night stayover requirements are used to prevent business travelers from taking lower fares. Thus, it appears we have sustainable price discrimination in a competitive industry.

But there are many aspects of the airline industry that indicate market power. Entry is limited by landing rights and gates. This makes it difficult for a rival to enter and attract away the profitable business customers. With the recent expansion of secondary airports, we are witnessing the entry of all-business discount airlines. Between New York to London, Maxjet, Silverjet, and Eos have all substantially undercut the business class fares on British Airways, American, Virgin, and Continental.

In other cases, undercutting does not occur because product differentiation gives the incumbent airlines a degree of market power. Some customers are not willing to take the lowest fare regardless of the airline or departure time. The flight variety offered out of an airline’s hub—for example, Northwest out of Minneapolis—may be a superior product to a rival’s single competing option. Or dedication to a frequent flyer program may be the source of differentiation.

Furthermore, airlines are involved in a complicated repeated interaction. A fare cut on one route may lead to retaliation on another. With limited entry, implicit cooperation allows firms an opportunity to exploit market power.

Thus, one can accept that there is some price discrimination in airline pricing and that it reflects limited entry, product differentiation, and repeated interaction—all departures from perfect competition. It is also important to note that the fact that airlines are not profitable does not contradict the existence of market power. The profits may be captured by pilots and a unionized labor force. This is a case of price discrimination but not an example of a perfectly competitive market.

Does this mean that price discrimination never arises in competitive markets? This conclusion might seem at odds with such everyday practices such as senior citizen discounts at movie theaters. Like the case of the local barber, the extent of price discrimination is quite small and is commensurate with a small degree of market power for a local movie theater.

provides fixed proportions of the two steaks. Thus, the true cost of producing more sirloin is the cost of rearing a cow net of the price of the flank steak. The true cost of producing more flank is the cost of rearing a cow net of the price of sirloin. Viewed in that light, the cost of the sirloin is higher than the cost of the flank steak.

*Airlines and movie cinemas share another unusual cost feature: Capacity is fixed, and marginal costs are zero until capacity is reached. Thus, the differential pricing reflects differential opportunity.*
The general point is that the extent of the price discrimination that a firm can impose is related to its market power. In truly competitive markets (which may be only a theoretical ideal), price discrimination is impossible. In real markets with product differentiation, the degree of market power is indicated by the markup of the tied good and the ability to exclude competitors. Evidence in the case indicated that Trident charged 2.5 to 4 times the price offered by Independent Ink. Competitors, such as Independent Ink, were excluded by the lifetime contract that was required with the print-head purchase. These features demonstrate a substantial and sustainable degree of price discrimination and hence market power.

Assume that market power has been established. This brings us back to the question of whether a firm should be allowed to engage in indirect metering via a tied sale of an essential complementary good. In the application of any per se rule, one has to be concerned about false positives: Legitimate applications of tied contracts will be prevented. Could what looks like a tied contract for metering actually serve a legitimate purpose, such as quality control or risk sharing?

ARE THERE LEGITIMATE REASONS FOR TIED CONTRACTS?

One has to be careful in stating the question. There is an important difference between offering the option of a tied contract and requiring a tied contract. In some cases, offering the option of a tied contract will enhance efficiency. But this will not explain why a customer is required to accept a tied contract.

Broadly speaking, there are three efficiency reasons used to justify a tying contract: (1) preservation of quality and reputation, (2) cost savings, and (3) risk allocation. However, quality control and cost savings should be achievable without contractually requiring the tie and do not explain the markup on the complementary good. The risk allocation justification can be accomplished with direct metering.

Preservation of Quality and Reputation

Firms may be motivated to tie for reasons related to quality and safety. For example, if a machine breaks down or the end result is unsatisfactory, the seller will suffer a loss of reputation. Thus, a firm may specify other inputs in order to ensure proper results.

---

While this motivation may apply in some circumstances, it has been invoked as an antitrust defense far more often than is likely to be present. The courts have usually rejected the quality explanation as a defense in tying cases, although the defense was accepted in the case of automobile spare parts.

In this case, Trident made the usual claim that Independent’s ink was of lower quality and caused damage to printheads, although Trident lacked evidence to support this finding. Chemical analysis showed that the inks were indistinguishable. In any event, industrial users should be able to judge the quality of the output or observe damage to the printheads. Moreover, even if the quality were inferior, the quality might be good enough and the price reduction significant enough that a user would choose to purchase Independent Ink’s product.

A basic problem with invoking the quality argument as an antitrust defense is that the compulsion of a tied sale should not be necessary to achieve the professed goal of ensuring quality. If the primary good manufacturer’s complementary or aftermarket products truly ensure quality, then customers should choose the firm’s products without being forced to do so. To the extent that Trident’s price differential reflected quality differences, few customers would be attracted to Independent Ink’s product. If the original product fails as a result of using inferior complementary or aftermarket products, the firm can help educate the customer that the failure was caused by other products and not by the original product.

A further problem is that the quality story would not lead Trident to charge a monopoly price for the replacement ink. If the pricing of its ink were comparable with that sold by Independent Ink, few customers would be willing to violate the contract and risk incompatible ink without the promise of substantial cost savings. Were Trident to be selling the ink at near competitive prices, it would not lose substantial profits when sales were diverted to Independent Ink. Thus, the only time the contract is required is when the complementary good is sold at a supracompetitive price.

**Cost Savings**

A related argument to quality is that the tied sale promotes cost reduction or enhanced efficiency. Evans and Salinger (2005, 2007) have identified many examples of such tying arrangements.

In some of these cases, the tied sale of complementary products can be viewed as one product rather than two. Thus, shoes are sold with shoelaces.
It is more efficient for the shoemaker to include laces of the right length with the shoes than to make the purchaser buy the initial pair of shoelaces separately.

The sale of a printer with the initial toner or ink cartridge is another such example. The least-cost way of delivering an initial cartridge is along with the printer. Selling printers with an initial cartridge is a contemporaneous tied sale. Unlike the sale of replacement cartridges, it is not a case of metering. All customers get the same number of initial cartridges, namely, one. Furthermore, such a tied sale is *de minimus* in that it affects only a small percentage of the total market. If the average customer purchases twenty cartridges over the lifetime of the printer, the initial sale covers only 5 percent of the market.

The efficiency or one-good arguments generally do not apply to the case of a requirements tie. No one imagines that a lifetime supply of shoelaces is really part of the original shoe and shoelace package. Nor do we imagine that there are cost savings associated with requiring the customer to purchase all future shoelace needs from a specified manufacturer or all of one's future ink requirements from Trident. To the extent that there are cost savings or economies of scale, they should be reflected in lower prices that lead consumers to purchase aftermarket products from the most efficient supplier.

Cost efficiencies thus are not a valid explanation for why a firm would require its customers to make all complementary good purchases from itself or a specified vendor. The efficiency defense of a requirements tie is less compelling than for a contemporaneous tie because there is little danger that dissolving the requirements tie would disrupt an efficient production or distribution arrangement.

**Risk Allocation**

Trident argued that a requirements tie is efficient because it improves risk allocation. When customers are uncertain about the value of the product, sellers can reduce the risk via a no-lose contract. The buyer pays a small amount up front and pays a price with each use, making the sale more like a lease. In effect, the customer pays for the product only if and when he or she uses it. This type of contract also signals the seller's confidence that the buyer will value the original good. The per-use sales contract is typically done via a requirements tie for a consumable product used with the original item.

Justice Breyer emphasized this point during the oral argument. He posed the following hypothetical:

"[T]he person is trying to establish the market for the product. It's a component, and he attaches this tied product as a counting device knowing that..."
if it's successful, everybody makes money, and if it's not successful, he and everybody else lose.

A seller should be able to accept the risk for an unproven technology. Metering is often the way by which this type of risk sharing arises. Since the world of patents is the world of unproven technologies, if metering were to lead to antitrust violations whenever there was a patent involved, this would create a chilling effect on inventors. While risk allocation can be an efficient use of tying, it does not appear relevant to this case. Trident's inkjet printheads are no longer new or unproven. If risk allocation were a motivation, then the customer would be asking for a metering contract—it would be an option, not a requirement imposed upon the customers. Customers would have the option of buying a printer at one price without a requirements tie or at a lower price with such a tie.

Even where risk allocation requires metering, it does not require the use of a tied sales contract. The metering can be done directly. There is nothing in the law that would prevent Trident from putting a counter in its printheads or requiring printer manufacturers to place such a counter in the printer. Trident is free to charge a price per copy or printout as part of the sale contract.

Historically, it might have been hard to monitor and therefore charge for direct usage. Today, with RFID chips and components connected to the Internet, this type of direct payment seems straightforward. Even without direct monitoring, the printer could stop working after each 1000 copies. The owner could then purchase a code for the next 1000 copies. It is worth emphasizing this point. The Court was not asked to outlaw metering or even price discrimination via metering. It was asked only to uphold the per se rule against a tied sale for metering where the metering was done indirectly via a separate product. Thus, efficiency loss arises only where risk sharing is desirable but direct metering is impractical.

This raises the question of why indirect metering should be outlawed while direct metering is permitted. The short answer is that metering should not be facilitated because it allows a firm to increase its exercise of market power. We return to answer this question at some length below. First we consider the traditional argument against tied sales to observe that they have little application to Trident v. Independent Ink.

“While direct metering is enhanced by technology, it appears to have been feasible in commercial applications for some time. In Morton Salt Co. v. G.S. Suppiger Co., 314 U.S. 488 (1942), the machines were used to add salt to canned food, and thus Morton could have measured how many cans of tuna were produced. As noted earlier, Monsanto, Summit, and Wizard International have implemented systems of direct per-use fees for their products. Trident could have the printer OEMs add a counter to their printers and thereby charge end customers a fee based on the number of cartoons printed.
Leveraging Market Power

In our discussion of tying, we have considered a range of motivations with the notable exception of leveraging market power. Yet, the fear of leveraging market power is the most commonly invoked justification for prohibiting tied contracts and the motivation behind most of the Court’s early decisions on tying.

It is also controversial. This leveraging motivation for tying has been subject to harsh criticism from academic economists, especially the Chicago School. According to the Chicago School argument, a firm that has a monopoly in good A has only one monopoly profit to earn and thus cannot gain by leveraging its power into a second market for good B.

The flaw in the Chicago School argument is that it takes a static or short-run perspective. Although it is true that the firm that engages in tying may not increase its present one-monopoly profits, that is not the point of the tie. The tied sales are designed to help preserve the original monopoly or to lead to monopolization of the aftermarket product; see Whinston (1990), Carlton and Waldman (2002), and Nalebuff (2004). The tying of one good to another may reverberate back into the market where the firm has power. If the tying arrangement leads to exit in the tied-good market, this in turn may reduce the possibility of entry in the original market.

For example, a likely entrant might come from a firm in the tied-good market. This scenario was central to the U.S. antitrust case against Microsoft. There, the PC operating system was the original market, and an Internet browser was the tied market. Microsoft’s tied sales made it more difficult for Netscape to compete in the browser market. Even if consumers were no worse off for the loss of competition in the browser market, they may have lost one of the few potential entrants who could have challenged Microsoft’s market power in the operating system market.

Tying may also protect monopoly power through the loss of competition in the tied-good market. This becomes a form of raising rivals’ costs. Entry is deterred if the tied good is an essential complement to the original product and is no longer available (or is available only at a much higher cost) to a potential entrant.

If the firm controls enough of the market for the complementary product, then it may attain a monopoly in that market as well. Depending on the cost of entry and minimal efficient scale in this market, the product A monopolist may then enjoy the fruits of two monopolies. The Chicago School argument assumes that the product B market remains competitive. But, as a result of the requirements tie, the B market may over time become more concentrated or even monopolized.

“See United States v. Microsoft Corp., 253 F.3d 34, 50 (D.C. Cir. 2001) (en banc). See also the discussion of this case by Daniel Rubinfeld in Case 20 in this volume.
Although these examples illustrate an important theoretical point, leveraging market power does not appear to be a motivation for the tie in Independent Ink. It is unlikely that Trident’s exclusion of Independent Ink and others will lead to the monopolization of the ink market or to increased costs for future entrants into the printhead business.

Metering not only is the most powerful explanation for the requirements tie in Independent Ink but also appears to be the best explanation for the tying arrangement in many of the tying cases decided by the Supreme Court. The courts have been focused on the effects of leveraging when they should have been focused on the effect of price discrimination. This is true for International Salt, Leitch, IBM, Carbice, and A.B. Dick.”

The common feature of these cases is that firms with a patented product conditioned the sale or lease of that product on the purchase of an essential complementary product. The Court indicated that the purpose of the tying practice in these cases was to extend the firm’s monopoly to the tied market. As critics of these decisions have argued, it does not seem reasonable that commodity products such as salt, tar, buttons, punch cards, dry ice, or mimeograph fluid would be monopolized. It is more reasonable that the primary purpose of the tying contracts was to engage in price discrimination via metering. In each of these cases, the intensity of use of the original product was directly related to the consumption of the aftermarket product.

While the price-discrimination argument is the most compelling explanation of many tying contracts, this argument has rarely been invoked in court to justify such conduct. According to Kaplow (1985, at 545), defendants had only twice argued that the tied product sale is used as a metering device, and then only in summary fashion:

This alternative explanation has not appeared in a Supreme Court brief in the past 63 years. In United Shoe Mach. Co. v. UnitedStates, 258 U.S. 451 (1922), counsel for the appellant-defendant allocated one clause of a single sentence in a 1,562-page brief to the proposition that the tied products might serve as a meter to measure the use of the tying product. In Henry v. A.B. Dick Co., 224 U.S. 1 (1912), counsel for the appellee-defendant similarly argued that the tied product operated as a meter capable of measuring the use of the tying product so as to facilitate the determination of an appropriate royalty.

Most of the early cases preceded the economic theory of metering, and thus it is not surprising that metering is not analyzed in those cases. Modern defendants may be reluctant to raise metering as a defense for fear that the

"See Kaplow (1985).

court will see price discrimination as evidence of market power, which would then lead to a per se violation. As the Supreme Court has observed: "If, as some economists have suggested, the purpose of a tie-in is often to facilitate price discrimination, such evidence would imply the existence of power that a free market would not tolerate."

THE CASE AGAINST TRIDENT

The argument against the tied sale by Trident is not that it will leverage market power to another market. Rather, the argument is that the price discrimination via the tied sales is itself a manifestation of enhanced market power. Although the firm is granted a patent as a reward for innovation, that reward is limited in terms of time and scope. A firm should not be allowed to increase the exercise of its power through indirect metering.

This very issue confronts the lower courts. The Supreme Court remanded the case back to the lower courts, where Independent Ink will be given an opportunity to demonstrate that Trident has market power, and Trident will be given a chance to rebut that claim. If Trident is found to have market power, then its tied sales contract will be per se illegal. Why should that be the case? There is no real danger that its tied sales contracts will monopolize the ink market.

Indeed, these issues raised by Independent Ink led Evans (2006) and Wright (2006) to propose that the per se rule against tied sales for a firm with market power should be overturned in favor of a rule of reason. Given the facts of this case as presented, Evans argues that under a rule of reason, the benefits from price discrimination would outweigh the costs. (Evans would go even further and suggests that since this conclusion generally holds, tying should be viewed as per se legal.)

This leads us to the question of whether price discrimination is a problem. There is consensus that price discrimination generally hurts consumers. That might be enough. According to Bork and echoed by the Supreme Court in two cases:

The legislative histories of the antitrust statutes, therefore, do not support any claim that Congress intended the courts to sacrifice consumer welfare to any other goal. The Sherman Act was clearly presented and debated as a consumer welfare prescription. (Bork 1995, p. 66.)

There is an alternative view that it is only total welfare that should count. Indeed, some interpret Bork’s use of "consumer welfare" to mean


Given the common belief that price discrimination improves overall efficiency, this would suggest that price discrimination should not be an antitrust problem. However, the result that price discrimination improves efficiency is true only for the artificial case of perfect price discrimination. This ideal case leads many to believe that more price discrimination improves efficiency. No such result follows.

**Imperfect Price Discrimination Is Not Efficient**

Trident asserted that "[m]ost tying arrangements are economically beneficial." Yet, few of the claimed benefits apply to the case of a requirements tie like that imposed by Trident. The predominant benefit claimed is that "patent tying may create efficiencies in calculating license fees" (at 29). An unstated, but implied, efficiency is the act of price discrimination.

In elementary economic textbooks, it is practically taken as a given that price discrimination is efficient, but this is an unfortunate oversimplification; see Carroll and Coates (1999). There are two problems with that conclusion.

The first problem is one of interpretation. The efficiency argument treats the impact on consumers and producers equally. The gain to the monopolist from price discrimination comes from increased sales and from an increased ability to capture surplus from consumers. Thus, while the monopolist gains, consumers generally lose. With perfect price discrimination, consumers end up with no surplus at all. Even if the monopolist gains more than the consumers lose, that net gain in efficiency should not justify allowing the tied contract. Producer profits and consumer surplus should not be (and are not) treated equivalently in antitrust.

The expected harm to consumers should result in striking down the tied contract even if there is a net gain in economic efficiency.

The second problem is fundamental: There is no reason to believe that price discrimination is efficient. The classic efficiency result is based on perfect price discrimination (also known as "first-degree price discrimination"). Perfect price discrimination requires the firm to have perfect knowledge about each customer's valuation. In practice, perfect price discrimination does not exist. Instead, price discrimination is based on

---

"Under special circumstances, price discrimination may result in a net gain to consumers. The price discrimination must lead to a large increase in output, and the increased output must result in significant economies of scale; see Hausman and MacKie-Mason (1988). This result also requires that the high-value consumers obtain very little surplus under the one-price-to-all contract.

"For example, §2 of the Sherman Act prohibits monopolization without regard to its efficiency. Thus, a firm charged with monopolization could not use the potential efficiency of price discrimination as a defense. Consumers are given more weight than producers when calculating damages in antitrust cases. Damages are not based on the loss of overall efficiency but rather on metrics such as overcharges or loss in consumer welfare."
purchase behavior such as the use of a complementary product (called "second-degree price discrimination") or on observable group characteristics such as senior citizen discounts (called "third-degree price discrimination"). In the simple case of linear demand and constant costs, both second- and third-degree price discrimination can result in lower efficiency. There is no general result that suggests that imperfect price discrimination improves efficiency, even treating consumer surplus and producer profits equally.

Of particular relevance to the Independent Ink case is the potential efficiency of price discrimination via metering tariffs. With a single consumer in the market and complete knowledge of the customer's preferences, a metering tariff can extract all the consumer surplus and achieve perfect efficiency. Bork (1995, p. 398) uses the intuition of perfect price discrimination to assert that "the more the monopolist is allowed to discriminate, the better are the results in terms of resource allocation."

This conclusion is false: With imperfect price discrimination, more is not always better. Specifically, the introduction of metering can reduce market efficiency when customers are heterogeneous. As the monopolist uses metering to capture more surplus from large customers, this may lead to small and medium customers' being excluded entirely from the market. As a result, total output and efficiency both may fall; see Kwoka (1984). There is no evidence, empirical or theoretical, suggesting that imperfect price discrimination in general and metering, in particular, improves efficiency.\footnote{Third-degree price discrimination causes a misallocation of goods to consumers, thereby reducing efficiency. Where demand functions are linear and the good is produced at a constant marginal cost, profits increase by less than the loss to consumers; see Schmalensee (1981).}

Furthermore, firms will find resistance to their attempts to impose imperfect price discrimination. This resistance leads to an often-overlooked inefficiency: Firms spend significant resources to impose price discrimination, while consumers spend resources to avoid being subject to price discrimination.

A third reason why price discrimination is costly is that no firm has a perfect monopoly. Consumers can find ways to create substitutes. They will do their best to avoid being subject to price discrimination, and these avoidance efforts end up being costly. Unlike the consumer welfare loss from a single-price monopolist, where low-value customers end up being excluded from the market, price discrimination causes high-value customers to distort their behavior, which can make the costs especially large.

Airline pricing illustrates this phenomenon. Airlines use advance purchase and Saturday night stayover restrictions to offer discounts to leisure

\footnote{A necessary, but not sufficient, condition is that price discrimination lead to an expansion in output; see Varian (1985).}
customers while still getting business travelers to pay full fare. In response to this policy, some business travelers extend their stays to include a Saturday night. This distortion in behavior can lead to large social losses, well above the incremental gains from the extra demand that arises at the discounted fares; see Morrison and Winston (1995) and Nalebuff (2003). The effort to price discriminate is also costly to firms that make large investments in "customer relationship management" (CRM) software in order to fine-tune their pricing to customers.

In sum, price discrimination via metering will typically lead to reductions in consumer welfare. There is no presumption that total welfare will increase. Even when there is an apparent gain, the efficiency calculations leave out the substantial costs that firms pay to enforce such metering and the substantial costs that consumers pay to avoid being metered. For these reasons, the courts should maintain the per se rule against tying, even when used for the purpose of metering.

**Metering, If Done, Should Be Done Directly**

Assume that it was established that the requirements tie in Trident was designed for the purpose of price discrimination via metering. Why should this metering be a problem when done indirectly via a requirements tie but not a problem if done directly? We accept that a legally gained monopolist is allowed to charge a monopoly price, even to the detriment of consumers. But the law need not facilitate the exercise of monopoly power. The argument against indirect metering is that it may be easier to implement because it is more opaque to consumers.

Firms anticipate consumer resistance to direct metering, such as a price per page. Metering via a requirements-tied sale is less obvious to consumers. With direct metering, the customer must be told what the cost is per page or per use. With indirect metering, customers may be in the dark. Market power is enhanced by the shrouded nature of the pricing.

Studies show that most customers have little idea of how much it costs to print a page using a laser printer or with inkjet cartridges; see Gabaix and Laibson (2006). Customers do not know the capacity of the ink or toner cartridge measured in typical pages. Even if they can forecast the future cost of the cartridge, they do not know the proper denominator. In the case of Trident, ink consumption rates were not provided, making it difficult to estimate the true lifetime cost of using its printhead.

Indirect metering via tied sales is thus a more subtle form of price discrimination. It achieves a result similar to direct metering without a blatant or naked exercise of market power. Thus, it may facilitate metering and thereby enhance a firm’s existing market power.

To the extent that risk sharing is valuable, it can still be done via direct metering. Indeed, to the extent that ink or the complementary product only
imperfectly measures usage, the firm can improve its profits through the use
of direct metering.

Firms that have been able to price discriminate only because cus­
tomers were fooled by shrouded prices would no longer be able to enhance
their market power at the expense of consumers. Such firms are using the
complementary product as a way artificially to enhance their market power.
Thus, the prohibition against tied sales for a firm with market power can be
justified as a way to limit the firm’s exercise of its market power through
price discrimination.

CONCLUSION

The predominant explanation for the requirements-tied sales contracts
used by Trident is price discrimination via metering. Customers were not
offered an opportunity to buy the printer at a higher price without the ink
contract. The contract went beyond the initial sale of ink to cover a life­
time supply. The replacement ink was sold at a very high price compared
with that of the alternatives. This ability to price discriminate via meter­
ing should be viewed as evidence of market power. Although the Court’s
ruling in this case means that Trident’s patent no longer leads to a pre­
sumption of market power, this may be established through traditional
economic analysis.

On the assumption that market power is established, the next debate in
antitrust is whether tying contracts used for the purpose of price discrimi­
nation should be viewed as a per se violation, subject to a rule of reason, or
presumed to be legal. Should metering be treated differently when done
directly versus indirectly?

I have argued that the per se rule should be maintained even when the
tie is used for metering. The rationale is explained by Justice Stevens in
Jefferson Parish:

[Tying] can increase the social costs of market power by facilitating price
discrimination, thereby increasing monopoly profits over what they would
be absent the tie. And from the standpoint of the consumer—whose inter­
est the statute was especially intended to serve—the freedom to select the
best bargain in the second market is impaired by his need to purchase the
tying product, and perhaps by an inability to evaluate the true cost of either
product when they are available only as a package.”

Whether one agrees with this perspective or not, the case of Trident v.
Independent Ink has set the stage for a much-needed reexamination of the
broader rule against tying.

REFERENCES


Carville Corp. of Am. v. American Patents Development Corp., 283 U.S. 27 (1931).


Motion Picture Patents Co. v. Universal Film Co., 243 U.S. 502 (1917).
INTRODUCTION

On January 5, 1999, the Antitrust Division of the U.S. Department of Justice (DOJ) filed a complaint alleging that Dentsply International, Inc. (Dentsply) had violated Sections 1 and 2 of the Sherman Act and Section 3 of the Clayton Act by using a variety of actions to maintain its monopoly power in the U.S. market for artificial teeth. Of central concern was Dentsply’s policy of refusing to sell its Trubyte brand of teeth to dealers carrying certain lines of competing artificial teeth. The complaint alleged that this policy harmed competition because it prevented rival artificial tooth manufacturers from obtaining adequate dealer networks. The Government alleged that consumers consequently suffered from higher prices, lower quality, and less choice.

On April 3, 2000, Dentsply moved for summary judgment, arguing inter alia that its policies could not have significant anticompetitive effect because there were other viable methods of distribution and because the contracts between Dentsply and its dealers were at will. Chief Judge Sue

*At the time this case went to trial, the author served as deputy assistant attorney general for economic analysis. This chapter draws solely on publicly available information, and the views expressed are not necessarily those of the U.S. Department of Justice.


Various documents, including briefs and expert reports, were filed by Dentsply under seal. Here and below, the descriptions of Dentsply’s arguments are primarily based on the plaintiff and/or
L. Robinson denied Dentsply’s Motion for Summary Judgment and held a bench trial (i.e., there was no jury) in April and May of 2002.

The judge issued her opinion on August 8, 2003. As has become customary, she issued a long, factually detailed opinion. She found that Dentsply had long had a very high market share, had set prices with little regard for competitor reactions, and had imposed its exclusive dealing policy with the express intent of limiting the ability of rival artificial tooth manufacturers to compete effectively. Nevertheless, the judge ruled that Dentsply had violated neither Section 1 nor Section 2 of the Sherman Act nor Section 3 of the Clayton Act because its exclusive dealing policies did not have anticompetitive effects. Specifically, she found that other modes of distribution—notably direct distribution from manufacturers to their customers—were viable and that Dentsply’s dealer agreements did not foreclose other manufacturers because the agreements were at will.

The DOJ appealed the district court’s decision solely with respect to the Section 2 claim. On May 26, 2005, a three-judge panel of the U.S. Court of Appeals for the Third Circuit reversed Judge Robinson, finding clear error with respect to several factual interpretations. The appellate panel concluded that dealers were bound to Dentsply by virtue of its market position (rather than formal contracts) and that, although alternative forms of distribution were “viable,” they were not sufficient to mitigate fully the adverse competitive effects of Dentsply’s policies.

After Dentsply’s petition for a writ of certiorari (i.e., a request to be heard on appeal) with the U.S. Supreme Court was denied, Judge Robinson (pursuant to the Third Circuit’s opinion) issued a final judgment on April 26, 2006, prohibiting Dentsply from requiring exclusive distribution of its artificial teeth.

Like many if not most rule-of-reason decisions, this one was very fact intensive. The case is, however, notable for the broader principles that it affirms. First, it applied the standard that exclusionary behavior that significantly raises rivals’ costs can harm competition to the point of constituting a Section 2 violation and does not have to constitute total foreclosure. Second, it rejected the position taken by several courts that short-term contracts could not give rise to anticompetitive effects.


INDUSTRY BACKGROUND

Dentsply supplies a wide range of dental products ("from alloys to x-rays") in countries around the world, including the United States. The focus of the case was on Dentsply's sale of artificial teeth in the United States through its Trubyte Division. Although that division also sold other dental merchandise, tooth sales accounted for approximately 80 percent of its revenues.

Artificial teeth are used in the manufacture of dentures and other restorative appliances. Because appearance and fit matter, artificial teeth are made in thousands of different shades and shapes. There are also several different grades (quality levels) of teeth. Dentsply manufactured over 100,000 different types of tooth units, and its Trubyte Division's teeth were in the premium and economy segments of the market. Trubyte's 2001 net tooth sales totaled $40.4 million.

At the time of trial, Dentsply had for many years had a high and stable market share: 75 to 80 percent on a revenue basis. As shown in Table 14-1, no other manufacturer had more than a tiny share of the market. In addition to those shown in Table 14-1, there were a handful of other manufacturers with even smaller sales of artificial teeth in the United States.

Although there had been little growth in the market for several years, entry had occurred. Heraeus Kulzer GmbH, a German artificial teeth manufacturer, entered the U.S. marketplace in January 2000 and introduced teeth specifically designed to meet U.S. preferences. Davis, Schottlander & Davis, Ltd., an English manufacturer, entered in 2001.

As illustrated in Figure 14-1, Dentsply sold its artificial teeth through independent dental product dealers, which sold the teeth to dental laboratories. The dental laboratories, in turn, fabricated dentures in response to requests from dentists acting on behalf of their patients. Notably only 10 percent of dentists specified the particular brand of teeth to be used, and labs generally chose the brand. When a dentist did specify the brand, however, the laboratory was obligated to use that brand.

At the time of trial, there were approximately forty thousand dentists in the United States who worked with dentures. There were approximately sixteen thousand dental laboratories that fabricated dental restorations, of
### TABLE 14-1
**Market Shares**

<table>
<thead>
<tr>
<th>Manufacturer</th>
<th>Percentage Market Share by Revenue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dentsply</td>
<td>75-80</td>
</tr>
<tr>
<td>Ivoclar Vivadent, Inc.</td>
<td>5</td>
</tr>
<tr>
<td>Vita Zahnfabrik</td>
<td>3</td>
</tr>
<tr>
<td>Myerson LLC</td>
<td>3</td>
</tr>
<tr>
<td>American Tooth Industries</td>
<td>2</td>
</tr>
<tr>
<td>Universal Dental Company</td>
<td>1-2</td>
</tr>
<tr>
<td>Heraeus Kulzer GmbH</td>
<td>1</td>
</tr>
<tr>
<td>Davis, Schottlander &amp; Davis, Ltd.</td>
<td>&lt;1</td>
</tr>
</tbody>
</table>

### FIGURE 14-1 Distribution of Artificial Teeth

- Manufacturer → Dealer → Laboratory → Dentist → Patient
which approximately seven thousand fabricated dentures. There were hun-
dreds of dealers, although the industry was consolidating, and several deal-
ers were much larger than others in both geographic scope and distribution
volume. Sales among Dentsply’s own dealers were quite unevenly distrib-
uted, with the two largest dealers accounting for 67 percent of total sales,
and the five largest accounting for 83 percent."

There were also different types of dealers in the overall dental supply
industry. Laboratory dealers served dental labs. In addition to artificial
teeth, these dealers offered a wide range of other materials used by dental
laboratories. At the time of trial, Dentsply sold approximately $400 million
of other products through its dealer network. "Operatory dealers “carr[ed]”
products for dentists exclusively or in combination with dental lab prod-
ucts.” The dealers shown in Figure 14-1 are laboratory dealers. When not
otherwise distinguished, in the text that follows "dealer" refers to labora-
tory dealers.

As Figure 14-1 illustrates, not all teeth went through distributors.
Several manufacturers—including Ivoclar and Heraeus Kulzer GmbH—
sold their teeth directly to dental laboratories. Dentsply itself had consid-
ered selling directly to the laboratories. According to Dentsply executives,
the company abandoned plans to do so because it feared that it would alien-
ate dental dealers, on whom Dentsply would continue to rely to distribute
its other dental products. Other manufacturers, such as Myerson L.L.C. and
American Tooth Industries, simultaneously sold artificial teeth directly and
through dealers, including Dentsply dealers. Lastly, Vita Zahnfabrik was a
German manufacturer that sold its artificial teeth in the United States
through its affiliate, Vident. Although Vident was the “exclusive dealer” for
Vita, Vident had arrangements with several other companies that served as
“subdealers” to distribute Vita teeth. For purposes of analysis, one might
better view Vident as an extension of the manufacturer rather than an exclu-
sive dealer.

Even companies that relied heavily on dealer networks frequently
bypassed their dealers in the physical distribution of artificial teeth. The
ability to provide teeth rapidly (especially when replacing existing pros-
theses on an emergency basis) was important to patients and, hence, to den-
tists, laboratories, and dealers. Laboratories tended to have inventories with
large numbers of teeth but had to replenish those inventories and also had to
order specific teeth that the labs did not have in stock but needed to fabri-
cate a particular denture. Market participants would often rely on drop ship-
ments (i.e., a dental lab would place an order with a dealer, but the teeth
would be shipped directly from the manufacturer to the dental lab). At the

is a typographic error in the opinion at p. 406; "non-Dentsply" should read "non-Trubyte."
time of trial, approximately 60 percent of orders for Dentsply’s Trubyte teeth were drop shipped.\textsuperscript{14}

Dentsply had long had an informal policy of discouraging its dealers from adding competing tooth lines to their distribution. Dentsply formalized its policy in 1993 with the adoption of Dealer Criterion 6: “In order to effectively promote Dentsply/York products, dealers that are recognized as authorized distributors may not add further tooth lines to their product offering.”\textsuperscript{15} Dealers were grandfathered if they had been carrying competing lines of teeth at the time Dealer Criterion 6 was adopted, and Dentsply’s five largest dealers in 2001 all sold competing lines of artificial teeth under grandfathering provisions. The grandfathering covered only lines that were offered at the time; the grandfathered dealers were not allowed to add new lines of competing teeth after 1993. Dentsply actively enforced Dealer Criterion 6 and conditioned distribution of Trubyte products other than teeth on whether a dealer carried non-Trubyte teeth.

SOME ECONOMICS OF EXCLUSIVE DEALING

The central economic issues in the case revolved around the effects of Dentsply’s exclusive dealing policy on competition and consumer welfare. At a very high level, the economic evaluation of exclusive dealing in a given case comes down to determining whether it (1) harms competition by denying rivals effective, low-cost access to significant numbers of potential customers and thus weakening rivals’ ability to compete or (2) promotes efficient distribution by aligning manufacturer and dealer interests.

Economists and lawyers have developed several theories regarding how exclusive dealing can give rise to effects of types 1 and 2.\textsuperscript{16} As will become apparent, the research literature has taken two significantly different approaches to analyzing the effects of exclusive dealing. In very broad terms, the old-school approach is what could be termed a partial partial-equilibrium analysis. This approach makes informal arguments based on a very limited set of contractual and organizational possibilities. It tends to take some features or parameters (e.g., the wholesale price) as exogenously given even when there is little reason to believe that such assumptions are appropriate.\textsuperscript{17} This approach has proven popular with the courts because it tends to generate simple and clear stories about competitive effects.

\textsuperscript{c}In the interest of space, this section reviews only the most prominent theories. For a discussion of additional theories, see Katz (1989).
\textsuperscript{d}This approach also tends to use rather loose proxies for welfare. For example, increases in promotional activities are often assumed to increase consumer and total surplus, even though there are circumstances under which this would not be the case.
A coarse description of the new-school approach is that it uses modern contract theory to conduct a complete partial-equilibrium analysis. That is, the analysis allows the parties to make use of a wide range of contracts and strategies, subject only to limitations imposed by the information structure and any legal prohibitions. Price levels and other elements of vertical relationships are determined as part of the equilibrium analysis. The cost of this approach is that it typically leads to complex, hard-to-apply results and finds very subtle linkages between behavior and welfare. It is thus not surprising that such theories appear to have had little impact on legal decision-making.

Before reviewing specific theories of harm and benefit, it is useful to address a sweeping argument taken by some to prove that exclusive dealing never poses a threat to competition:

A seller who wants exclusivity must give the buyer something for it. If he gives a lower price, the reason must be that the seller expects the arrangement to create efficiencies that justify the lower price."

Proponents of this argument failed to observe that it relies on multiple, often-unrealistic assumptions. One is that final consumers are either indifferent to the outcome of manufacturer-dealer negotiations regarding exclusive dealing or are party to those negotiations. Another is that either dealers act in concert or there are "pivotal" dealers that can unilaterally determine whether exclusive dealing succeeds in excluding rivals. When there are multiple dealers, one dealer's decision to enter an exclusive arrangement can—by contributing to the foreclosure of other manufacturers—have negative effects on the other dealers, as well as on the ultimate consumers. When dealers neither coordinate among themselves nor can easily coordinate with alternative manufacturers, any given dealer may require little inducement to get it to accept a manufacturer's exclusive dealing proposal. Specifically, the dealer may reason that its decision will not affect the competitive alternatives that are available in the marketplace (i.e., the dealer is not pivotal), so the dealer may as well accept the proposal."

Strikingly, as pointed out by Rasmussen et al. (1991, p. 1143), a manufacturer need not have market power prior to the imposition of exclusive dealing in order to be able profitably to sign most or all dealers to exclusive contracts. Instead, what is needed is that each dealer expects that other dealers will sign exclusive agreements with that manufacturer. In theory,


This common intuition was summarized in Katz (1995, p. 708). It was formalized and more fully explored by Rasmussen et al. (1991). Segal and Whinston (2000) corrected certain inconsistencies in Rasmussen et al.'s analysis and showed that their conclusions apply with greater force when the seller can price discriminate and bargain sequentially with customers.
dealer expectations could focus on one of a set of otherwise identical manufacturers. More plausibly, dealers will base their expectations on some underlying asymmetry among the manufacturers. Time is one possible dimension of asymmetry: An incumbent manufacturer may be able to enter into exclusive dealing arrangements before potential entrants are in a position to bargain with dealers. "As discussed below, there are possible asymmetries.

In assessing the effects of exclusive dealing, it is important to have a baseline of comparison (a "but-for world"). A natural benchmark is to consider situations in which two or more manufacturers use a single dealer, or common agent. Although the idea runs counter to the intuition that exclusive dealing can harm competition, economic theory has identified the use of a common dealer as a means of facilitating collusion. Interbrand competition can be viewed as a negative externality across manufacturers. One way to internalize this externality is to vest all of the decision-making power in a single economic actor who has been made the residual claimant to industry profits. If it is risk neutral, this residual claimant will maximize industry profits. Bernheim and Whinston (1985 and 1998) identified a common dealer as potentially serving this role." If all manufacturers sold their output to a single downstream firm, and each manufacturer set its wholesale price equal to marginal cost, the common dealer would act to maximize the collective profits of the two stages, which could be shared through fixed, lump sum transfers.

Given these benefits of common agency, why would manufacturers ever choose to have exclusive dealers? There are two broad answers. One is that, when there are more than two manufacturers or more than one potential dealer, it might be possible for one manufacturer to hire an independent dealer and free-ride on any output or promotion restrictions undertaken by the other, cartelized brands. As yet, this situation has not been fully modeled. The other possibility is that there are many circumstances in which it is not privately optimal for a manufacturer to set its wholesale price equal to marginal cost. Indeed, there can be foreclosure even when there is only one dealer, which is thus pivotal. One might expect this dealer to be well positioned to demand an efficient contract. But by signing a long-term exclusive dealing agreement with penalty clauses, the dealer and the incumbent manufacturer can force a manufacturer that later enters the market to compensate the dealer for breaking the agreement. In this way, the dealer and incumbent manufacturer can appropriate some of the benefits from entry in the form of higher profits for themselves (Ashenfelter and Bolton 1987). Here the externality is imposed on the entrant. For a related theory, see Bernheim and Whinston (1998, Section IV).

In a related vein, Yong (1999) observed that, when dealers compete with one another in the downstream product market, there are circumstances in which an entering manufacturer can—by playing the dealers off against one another—appropriate greater rents than it creates for the industry overall. Thus, even socially efficient entry can be undesirable from the collective perspective of the incumbent manufacturer and dealers.

"For a related argument with respect to tying, see Brennan and Kimmel (1986)."
to marginal cost, and thus meaningful agency problems arise. The next two arguments consider the consequences.

**Improving Dealer Focus**

According to the dealer-focus rationale, a multiline dealer will fail to provide high levels of promotion and service for any one manufacturer because it will also be interested in earning profits from the sales of rival manufacturers' products. This is an old-school argument in the sense that it is not based on careful and complete modeling of the vertical contracting problem and the nature of interbrand competition. The common agency model shows just how limited the dealer-focus intuition can be—under some conditions, common agency leads to the dealer's undertaking the joint-profit-maximizing levels of promotional activities.

A modern way of describing the dealer-focus motivation is to state that there are conditions under which having an exclusive agent reduces a principal's costs of inducing a desired level of effort from its agent. Martimort (1996) provides a rigorous model of this effect and shows that it can affect the distributor's equilibrium choice of actions. Martimort (1996, p. 16) explains the effects of exclusive dealing as follows:

Common agency differs from exclusive dealing in two respects. First, under common agency, a single agent takes both decisions on the downstream market. Even though a high level of coordination is then feasible, some inefficiency is introduced by the opportunity for this agent to misreport to two different principals. Under exclusive dealing, an agent is concerned only by the downstream profit on one product. This results in a lack of coordination. However, at the same time each agent can misrepresent to only one principal.

Martimort shows that, under some market conditions, the agency considerations dominate the coordination considerations, so that the private parties choose exclusive dealing. This model thus supports the dealer-focus argument.

There is another important implication of this model. There is a tradition of asserting that, if dealers oppose exclusive dealing, then one should

---

22 For a general discussion in the context of vertical contractual relations, see Katz (1989).
23 There is also another problem with old-school arguments such as dealer focus. When there are multiple manufacturers, industry profits may be greater if dealers undertake less promotional activity than that which would maximize the profits of any one manufacturer-dealer pair taking the activities of rival manufacturers as given. This fact raises several questions about whether manufacturers will seek exclusive dealing. See Besanko and Perry (1993) for a formal model.
24 One cost of this rigor is that the model is solved for the case of only two manufacturers, each of which can utilize at most one dealer.
25 See also Bernheim and Whinston (1998).
be concerned that it is inefficient. This line of reasoning is incorrect. Dealer opposition may simply reflect a belief that dealer rents would be larger under an alternative arrangement. Martimort (1996) establishes conditions under which exclusive dealing leads to greater output, consumer surplus, and manufacturer profits than does common agency. The manufacturers favor exclusive dealing in this case—and the dealers oppose it—because it allows the manufacturers to extract more rents from their dealers than would common agency. Here manufacturer interests are aligned with consumer interests and efficiency, while dealer interests are opposed.

Protecting Manufacturer Investments

The dealer-focus theory concerns the effects of vertical contracting on competitive activities undertaken by dealers. The next two theories concern the effects of vertical contracting on competitive activities undertaken by manufacturers.

Under the logic put forth by Klein (1988), exclusivity promotes manufacturer investment in relationship-specific assets by enabling a manufacturer to extract a greater share of the available surplus in bargaining with its dealer because the latter cannot hold up the manufacturer by threatening to transact with another manufacturer. Observe that this intuitive analysis applies to situations in which the investments have no value outside of the specific manufacturer-dealer relationship, as would be the case, for example, if a manufacturer trained a dealer how to use specialized data systems that were proprietary to that manufacturer.

The next theory applies only to situations in which a manufacturer’s investment increases the value that a dealer could earn in its relationship with the investing manufacturer and with one or more other manufacturers. An example would be a manufacturer’s providing a dealer with generalized training that allowed the dealer to sell any brand more effectively. When the dealer applies that training, it has incentives to persuade consumers to purchase whichever brand offers the largest profit margin. That brand need not be the one that invested in training the dealer. Foreseeing this free-riding across manufacturers, manufacturers may invest less than what would maximize the sum of manufacturer and dealer profits.\(^26\) Marvel (1982) identified the imposition of exclusive dealerships as one way for a manufacturer to prevent its dealer from engaging in cross-brand steering.\(^27\)

Both Klein and Marvel made old-school arguments. In contrast, Segal and Whinston (2000) took a new-school approach to analyzing the investment incentive effects of exclusive dealing in a formal model. They found

\(^{26}\) Observe, however, that, in an oligopoly, free-riding might increase manufacturer profits by inducing promotional efforts closer to the collusive level.

\(^{27}\) At one level, this effect is an extension of the dealer-focus theory.
that exclusivity has no effect if the manufacturer’s investments do not influence its dealer’s value of transacting with other manufacturers. Specifically, they show that, when manufacturer investment does not affect the dealer’s value of distributing other manufacturers’ products, exclusive dealing increases the manufacturer’s share of overall surplus but does not increase the marginal returns to manufacturer investment under reasonable assumptions about the bargaining process. Because marginal returns are what matter for investment incentives, this analysis strongly suggests that Klein’s argument is incorrect. However, Segal and Whinston also found that exclusivity does matter when a manufacturer’s investments do affect the dealer’s value of transacting with other manufacturers, as occurs under Marvel’s free-riding theory.

Foreclosing Rivals from Doing Business

Having looked at theories of how exclusive dealing might promote efficiency, now consider two theories of how it might harm competition. First, an old-school analysis focused on whether rival manufacturers were literally foreclosed from serving consumers, as would happen if most or all dealers were tied to a manufacturer or if final consumers signed exclusive contracts directly with the manufacturer. This analysis emphasized the importance of contract length in markets in which temporal asymmetries provide the mechanism by which a manufacturer is able to obtain exclusive dealing agreements at relatively low cost. Specifically, if all contracts were of short duration, then a new manufacturer could simply wait for them to expire and quickly obtain distribution comparable with that of the incumbent. But if an incumbent had long-term, staggered exclusive dealing contracts with the best dealers, an entrant will be unable to obtain equivalent or necessary distribution without a considerable lag (unless it induces breach).

Time, however, is not the only dimension of possible manufacturer asymmetry. One manufacturer may—because of lower production costs or more desirable product characteristics—have a much higher market share than its current rivals. If a dealer has to go with a single brand, it will tend to choose the market leader ceteris paribus. Hence, a manufacturer with a large market share may be able to tie up dealers even with at-will contracts.

One way to think about the traditional foreclosure theory is that exclusive dealing arrangements increase rival manufacturers’ distribution costs by so much that these firms are unable to compete. Economic analysis clearly establishes that consumer welfare and economic efficiency can be harmed by a practice that raises rivals’ costs even if it does not drive them out of business entirely. The next theory considers the use of exclusive dealing as such a practice.

See, e.g., Marvel (1982, p. 6).

For a broader discussion of cost-raising strategies, see Salop and Scheffman (1987).
Raising Rivals' Costs

A manufacturer's profits typically are an increasing function of its rivals' costs. Exclusive dealing arrangements can serve as a means of raising those costs. Suppose there are economies of scale and scope in distribution. Then a system of exclusive dealers will raise the distribution costs of all manufacturers but will do so more for the smaller ones than the larger ones. The net effect may be to raise the profits of larger manufacturers, even though their costs of distribution are raised as well.

Three observations about this theory are useful. First, a manufacturer does not need to enter exclusive dealing arrangements with all or even most dealers in the industry for this policy to be effective. Even if there are unsigned dealers available to serve rival manufacturers, these dealers will not be able to realize economies of scale and scope. Moreover, the effects of exclusive dealing can be amplified to the extent that dealers are heterogeneous, and the most efficient or effective dealers are tied up through exclusive dealing. Second, contractual lock-in is not needed for exclusive dealing to raise rivals' costs. Even at-will contracts can have adverse effects on competition and consumer welfare. Third, the welfare effects of raising rivals' costs are potentially complex. In many instances, actions that raise rivals' costs without lowering another firm's costs or raising its product quality will decrease economic welfare by increasing quality-adjusted prices. There are, however, circumstances under which such actions can increase total surplus or the welfare of some or all consumers. A fact-specific inquiry is needed.

THE PLAINTIFF’S ARGUMENT

In its complaint and at trial, the DOJ alleged that Dentsply’s exclusive dealing arrangements constituted unlawful agreements in restraint of interstate trade in violation of Section 1 of the Sherman Act; were an unlawful action to maintain a monopoly in violation of Section 2 of the

30 Indeed, Katz and Rosen (1985) and Seade (1985) showed that when marginal costs are increased by some action, even a symmetric (across all manufacturers) cost increase may raise a manufacturer's profits.

3 For example, raising an inefficient supplier’s costs can increase total surplus by shifting share to more efficient producers (see Katz and Shapiro 1985). Or increased costs of certain actions might force rivals to alter their distribution strategies in ways that benefited some consumers. Lastly, the elimination of a rival and the consequent reduction in product-market competition might promote R&D and other investment—an argument often associated with Joseph Schumpeter. For discussions of the literature examining the effects of increased firm size and concentration on innovation, see Gilbert (2007) and Katz and Shelanski (2007).

Sherman Act; and were restrictive dealing agreements prohibited by Section 3 of the Clayton Act.

The Government asserted that the relevant market was "the sale of pre-fabricated, artificial teeth in the United States" and that Dentsply had held monopoly power in this market at least since 1987. The Government pointed to, among other things, Dentsply's high and relatively stable market share as evidence. The DOJ also argued that Dentsply's margins for artificial teeth were high relative to its margins on other dental products and rival manufacturers' margins on artificial teeth and that those high margins were consistent with the possession of monopoly power.

The Government alleged that Dentsply's exclusive dealing policies limited the ability of other manufacturers to obtain effective and efficient distribution and thus weakened—but did not eliminate—competition in the sale of artificial teeth. This allegation rested on two fundamental claims. First, the Government claimed that Dentsply's policies denied other tooth manufacturers access to the dealers aligned with Dentsply. The Government did not assert that Dentsply locked in dealers with long-term contracts. Instead, the Government alleged that, when faced with an all-or-nothing choice, a dealer would choose to distribute Trubyte products rather than those of one or more of Dentsply's much smaller rivals. The Government also asserted that Dentsply's practice of conditioning distribution of Trubyte products other than teeth on whether a dealer carried non-Trubyte teeth increased a dealer's cost of carrying rival tooth brands.

Second, the Government's allegation rested on the claim that competing tooth manufacturers did not have available to them close substitutes for dealers aligned with Dentsply. The Government asserted that Dentsply had tied up the most important and valuable dealers, that operatory dealers were not close substitutes for laboratory dealers, and that direct distribution was not an effective substitute. According to the Government, Dentsply's dealer network constituted approximately 80 percent of laboratory dealers. Although recognizing that manufacturers had the ability to engage in direct distribution, the Government argued that dealers provided several benefits to tooth manufacturers, either directly or indirectly. The direct benefits identified by the Government included holding inventory, handling accounts receivable, and providing additional sales representatives and advertising channels. The Government also identified several benefits to dental labs and asserted that these lab benefits indirectly benefited tooth manufacturers. The indirect benefits included local availability of teeth and "one-stop" shopping offered by dealers to laboratories to meet their artificial tooth and other dental supplies needs. Observe that the DOJ's argument depended on Dentsply's dealers' being better able to provide these benefits than were other dealers or the tooth manufacturers themselves.

*Plaintiff's Proposed Findings, § IV.B. 2.

Page 401
According to the DOJ, in addition to its direct effects on the strength of competition,

"[t]he lack of a substantial dealer network . . . has led [other tooth manufacturers] to curtail their promotional efforts, because doing so would be futile without an adequate dealer network."

The Government argued that the effects of lessened competition were higher prices, lower product quality, fewer consumer choices, and reduced flows of information to laboratories regarding product characteristics. The Government also asserted that these adverse effects flowed through to dentists and their patients.

**DENTSPLY’S DEFENSE**

Dentsply defended its exclusive dealing policy on several grounds. First, Dentsply argued that rival tooth manufacturers were not—and could not be—foreclosed because they could and did employ alternative distribution methods, including direct supply of dental labs through mail order. Dentsply argued that, because its policies did not completely foreclose competition from rival tooth manufacturers, the policies were not illegal. In making this argument, Dentsply implicitly asserted that the raising rivals’ costs argument identified above could not serve as the basis for finding an antitrust violation. Dentsply also asserted that rivals’ low market shares were not evidence of the effects of successful exclusionary behavior but were instead the result of choices that rival manufacturers had made.

Dentsply and its expert argued that, because dealers were not bound to it by express long-term contracts and were free to cease carrying Dentsply’s teeth at any time in order to carry rival brands, its policy was neither illegal nor anticompetitive. Here Dentsply rejected economic theories based on asymmetries other than temporal ones. Dentsply did, however, address the issue of asymmetric market shares by making a "small-fish-in-a-big-pond

*Complaint, f 38.*

"In addition to the issues discussed below, Dentsply disputed many of the DOJ’s assertions regarding facts and their interpretation.

"There was some dispute with regard to the position taken by Dentsply’s expert on this issue. According to the DOJ, the expert first insisted that a necessary condition for exclusive dealing to be harmful was that it "foreclose the only way to get a product to consumers," but later "developed a theory that exclusive dealing is not harmful as long as rivals have access to means of distribution that are at least ‘remotely competitive’" (Plaintiff’s Proposed Findings, f 285(i)).

"There was also dispute with regard to the position taken by Dentsply’s expert on this issue. According to the DOJ, the expert at first insisted that exclusive dealing could cause harm only by preventing entry but later said that it could also cause harm by limiting the expansion of existing rivals (Plaintiff’s Proposed Findings, ¶ 285(i)).
that those Dentsply dealers with very low shares of Dentsply's tooth sales might actually increase their artificial tooth sales by carrying one or more competing tooth brands instead.

In addition to arguing that its practices did not harm competition, Dentsply and its economic expert offered two affirmative rationales or business justifications for its exclusive dealing policies. First, Dentsply asserted that it wanted to "focus" its dealers on selling its teeth. Interestingly, Dentsply's economic expert did not support this rationale, apparently relying on arguments put forth in Marvel (1982). That paper rejects the view that "exclusive dealing is designed by manufacturers to elicit dealer services" (p. 3) on the grounds that

a manufacturer can obtain additional dealer promotional services only by paying for those services directly, and the opportunity to do so is independent of whether the dealer handles one or several lines, (p. 5)

However, as shown by Martimort (1996), this claim is incorrect. The principal-agent problem faced by a manufacturer that is trying to motivate its dealer can be affected by whether the dealer also distributes products for other manufacturers.

Instead of offering the dealer-focus rationale, Dentsply's economic expert testified that the exclusivity arrangements promoted efficiency by supporting manufacturer investments (e.g., promotional activities) and policies (e.g., allowing tooth returns) that would otherwise be subject to free-riding. Here, too, the argument followed the logic of Marvel (1982). Free-riding would take place if a lab came to a dealer based on Dentsply's promotional or other activities, and the dealer then steered the lab to another, presumably higher margin, line of teeth.

DISTRICT COURT'S OPINION

Judge Robinson applied a rule-of-reason balancing approach to all three counts in the complaint. The court described the standards for the Sherman Section 1 and Clayton Section 3 claims as follows:

To establish a claim under § 1 of the Sherman Act, plaintiffs must show that: 1) there was a contract, combination, or conspiracy; 2) that unreasonably restrained trade; and 3) affected interstate commerce."

"Dentsply did not assign exclusive territories, and many of its dealers thus competed against one another.

For § 3 cases, the Supreme Court has noted that ""[i]n practical application, even though a contract is found to be an exclusive-dealing arrangement, it does not violate the section unless the court believes it probable that performance of the contract will foreclose competition in a substantial share of the line of commerce affected.""

The court characterized the plaintiff's burden as that of showing the existence of actual anticompetitive effects or proving that the defendant had "market power," here meaning the ability to raise prices above those that would prevail in a competitive market. With respect to the Section 3 claim in particular, the court applied a standard that asked whether rival tooth manufacturers literally were blocked from selling to dental laboratories rather than whether rivals were weakened as competitors.

With respect to the Sherman Section 2 count, the judge applied two standards. First, she asserted that Section 3 of the Clayton Act is more restrictive than Section 2 of the Sherman Act and thus that a finding that Dentsply did not violate Section 3 of the Clayton Act would imply that Dentsply did not violate Section 2 of the Sherman Act either. The judge also stated the usual standard for a stand-alone Section 2 analysis:

"The offense of monopoly under § 2 of the Sherman Act has two elements: (1) the possession of monopoly power in the relevant market and (2) the willful acquisition or maintenance of that power as distinguished from growth or development as a consequence of a superior product, business acumen, or historic accident." Eastman Kodak, 504 U.S. at 480 (quoting U.S. v. Grinnell Corp., 384 U.S. 563, 570-571 (1966)).

Relevant Market and Market Power

The district court found that the relevant market was "the sale of prefabricated artificial teeth in the United States" and that laboratory dealers were the "ultimate consumers" in this market because they generally selected the tooth brand.

The court adopted a standard legal definition of, and test for, monopoly power:

"Monopoly power is generally defined as the power to control prices or to exclude competition, and the size of market share is a primary determinant of whether monopoly power exists."


Ibid. (internal citations omitted).
Looking at Dentsply’s persistently high market share, the court concluded that “[s]uch a predominant market share is sufficient for the court to infer monopoly power.” However, the court also noted the need for further analysis and stated that “[t]he DOJ must also prove Dentsply has the power to control prices or exclude competition.” The court concluded that the Government had proved neither.

With respect to whether Dentsply controlled prices, the court rejected the DOJ’s evidence with respect to profit margins and, in any event, asserted that “high margins are to be expected in a market in which significant pré-sale promotion is employed.” The court took note of the evidence that Dentsply tended to ignore competition and was not under pressure to react rapidly to market developments. Oddly, the court concluded that

Contrary to an indication of control over prices and exclusion of competitors, any lack of urgency on the part of Dentsply to react to market demands is a result of a lack of competition pushing Dentsply to compete.

. . . this lack of competition is due to Vident’s and Ivoclar’s own business decisions, not Dentsply’s exclusionary practices.”

The court found that, because direct selling was viable, other dealers were available, and Dentsply’s own dealers could be drawn away, Dentsply did not have the power to exclude competitors from reaching dental laboratories. As part of its consideration of whether competitors were excluded, the court observed that entry and competitor expansion had, in fact, occurred. The court refused to take the poor performance of Vident and Ivoclar as evidence of successful exclusion. Instead, the court found that the low market shares of Dentsply’s rivals were the result of choices made by those rivals, including decisions to offer teeth that did not meet U.S. demands and limited promotional efforts compared with Dentsply’s.

In contrast to the judicial definition of monopoly power, economists frequently define monopoly power along the lines of a large amount of durable market power, where market power is defined by the firm’s facing a downward-sloping (rather than flat) firm-specific demand curve, so that maintaining price above marginal cost is profitable. For monopoly power to exist as defined by economics, something must prevent entry and expansion

43 Ibid. (internal citations omitted).

44 Ibid, (internal citations omitted).


46 Ibid.

47 Such a definition is implicit in the following statement by the DOJ:

In United States v. Brown University, 5 F.3d 658, 668 (3d Cir. 1993), . . . (“Market power, the ability to raise prices above those that would prevail in a competitive market. . . .”) . . . Monopoly power represents a higher degree of market power than is required under Section 1.
by rival firms and thus make it profitable for the monopolist to charge supra-competitive prices. But that something need not be the specific practice or practices at issue in the case, nor does there have to be complete elimination of competition.

In the instant case, the district court’s analysis of monopoly power focused on whether the specific practice challenged by the plaintiff excluded rivals.” As just quoted, the district court concluded that Dentsply’s rivals were weak competitors as the result of their own choices, rather than as victims of Dentsply’s exclusionary policies. If one believes that rival firms were behaving rationally, this distinction appears to be irrelevant to the assessment of whether Dentsply possessed monopoly power, at least as defined by economics. For whatever underlying reason, Dentsply faced weak and limited competition, and there was little reason to believe that the situation would soon change significantly (at least absent a change in Dentsply’s exclusive distribution policies). Thus, as a matter of logic, the district court’s finding that rivals performed poorly because of their own decisions was very likely not a sound basis for concluding that Dentsply lacked monopoly power.

Dentsply’s Efficiency Defense

Dentsply had argued that its exclusive dealing policy focused dealer efforts and prevented dealer free-riding on Dentsply investments. The district court found Dentsply’s efficiency defense to be pretextual. The court quoted a Dentsply document regarding the purposes of Dealer Criterion 6:

• Block competitive distribution points. Do not allow competition to achieve toeholds in dealers.
• Tie up dealers.
• Do not “free up” key players.”

The court offered additional theoretical and empirical grounds for rejecting Dentsply’s claim that dealer focus induced superior levels of promotion and service for Trubyte teeth. The court noted that grandfathered


Additional issues arise with respect to the definition of a competitive price when average cost exceeds marginal costs, and some economists define a competitive price as equaling average cost.

“The court did observe that entry had occurred, however, which implied that there was no factor that completely excluded rivals.

dealers had not been found to be less effective and cited testimony from dealers saying that focus effects would not arise. The court cited Marvel (1982) as rejecting this theory as a general matter. The court also offered its own theory tailored to the specifics of the case:

The "focus dealer services" rationale is not a valid justification for using exclusive dealing in the tooth industry because dealers have every incentive on their own to make sure that their level of service for any given tooth brand does not suffer. If a dealer provides inadequate service, it risks losing customers, not only for teeth, but also all the other products the customer may buy from the dealer. In fact, there is much greater risk to the dealer than to Dentsply. If a customer is dissatisfied with the service it receives from one Dentsply dealer, it will simply buy Trubyte teeth from another dealer. Dentsply’s sales will stay the same but the dealer will lose that customer’s business altogether.

Implicit in this rationale is a very strong empirical claim. The judge identified two sources of divergence between a dealer’s incentives to provide service and Dentsply’s economic interest in the dealer’s providing service. First, the dealer ignores the harm that poor service might cost Dentsply in terms of lost sales, which tends to lead the dealer to provide too little service from Dentsply’s perspective. Second, if a lack of service induces a customer to seek another Trubyte dealer, the underperforming dealing counts those sales as lost, even though from Dentsply’s perspective they are not. This leads to what Dentsply would consider excessive dealer incentives. The quotation above is a claim that these effects exactly cancelled one another. Although such a coincidence is a theoretical possibility, the decision cites no evidence that it was a reality.

The judge also offered several reasons for rejecting the claim that Dealer Criterion 6 promoted competition by preventing dealer free-riding. First, she found that this rationale was inconsistent with Dentsply’s stated reasons for adopting the policy. Second, she said that the defense had failed to provide factual support for its theory. The judge pointed to several potential empirical tests that the expert had chosen not to pursue or where the expert had looked but did not find evidence of free-riding. For example, the judge stated that the expert had produced neither any credible examples of free-riding by grandfathered dealers nor evidence that these dealers were in any other way less effective for Dentsply. Indeed, the judge rejected the possibility of free-riding on several grounds, including that dealers did not determine what brand of teeth were used in a particular denture case; dealers would be reluctant to alienate laboratory customers by recommending other brands of artificial teeth; and that much of Dentsply’s promotion was "purely brand specific."

The judge cited examples in which Dealer Criterion 6 was enforced in circumstances inconsistent with the claim that it was used to prevent free-riding. For instance, Dentsply terminated one dealer for adding a rival tooth line even though the dealer sold Trubyte merchandise but not Trubyte teeth.

The court also drew conclusions from the behavior of other parties. First, the court took the fact that dealers opposed Dentsply’s exclusive dealing provisions and some dealers even helped rival dealers evade the policy as evidence that the policy was inefficient.51 Second, the judge observed that no other artificial tooth manufacturer imposed exclusive dealing.52

Lastly, the judge observed that Dentsply’s own executives identified conditions under which the firm would increase its promotional activity in the absence of its exclusive dealing policy. The judge also asserted that, because of the increased effectiveness of their efforts, “. . . Dentsply’s rivals would increase their promotional expenditures if their teeth were sold through the dental laboratory dealer network.”53 It is notable that this last finding contradicts the court’s theory that Dentsply’s exclusivity policies were not responsible for the poor performance of rival tooth manufacturers.

Effects of the Contracts

Despite finding that “… Dentsply’s high market share could lead to a conclusion that Dentsply possesses monopoly power”54 and that Dealer Criterion 6 was intended to harm competition, the district court found "that Dentsply’s Dealer Criterion 6 is not an unreasonable restraint on competition.”55 The judge found that Dealer Criterion 6 did not block entry and competition because there were dealers outside of the Dentsply network available to distribute competing lines of artificial teeth; direct sales by a manufacturer to laboratories were a viable distribution strategy; and Dentsply’s dealers had at-will contracts and, consequently, were not locked in.

The DOJ argued that, as a practical matter, dealers could not leave Dentsply and, in fact none had. The judge did not accept this line of reasoning and instead found that dealers did not switch because Dentsply’s competitors had failed to compete. To the court’s way of thinking, the critical factor was that dealers could move away from Dentsply at will. The court

51 As discussed above, the logic of this argument is unsound.
52 Of course, other manufacturers also engaged in less promotion. Moreover, those manufacturers that solely engaged in direct selling would have no need for exclusivity policies (and could, indeed, be viewed as effectively having exclusive dealers).
Omega Envt'l, Inc. v. Gibarco, Inc., 127 F.3d 1157 (9th Cir. 1997) at 1163-1164, which stated that

[the] short duration and easy terminability of these agreements negate substantially their potential to foreclose competition. Because all of Gibarco's distributors are available within one year . . . a competing manufacturer need only offer a better product or a better deal to acquire their services. Plaintiffs' expert opined that no distributor would abandon the Gibarco line for an untested product with no reputation. We agree with the unremarkable proposition that a competitor with a proven product and strong reputation is likely to enjoy success in the marketplace, but reject the notion that this is anticompetitive, (internal citations and footnotes omitted)

Judge Robinson also rejected the Government's contentions that dealers provided several advantages over direct distribution and thus that foreclosure from dealers raised rivals' costs. In several instances, the judge rejected the DOJ's claims about specific functions and asserted that manufacturers could undertake the function themselves. In keeping with its approach of asking only whether other tooth manufacturers were completely foreclosed from competing with Dentsply, the court tended to ignore questions of whether manufacturers could offer services of as high or higher qualities at as low or lower costs as could dealers.

THE APPEAL

The DOJ chose not to appeal the district court's decision with respect to the Section 1 and Section 3 claims. Instead, the DOJ focused solely on the Section 2 claim and appealed on three issues, two of which were economic. The first was

Whether a firm that maintained a 75%-80% market share for a decade, established a price umbrella, successfully made repeated aggressive price increases without regard to the prices of its rivals, and was able to exclude rivals from a major channel of distribution, can be found not to possess monopoly power within the meaning of Section 2 of the Sherman Act, 15 U.S.C. 2, on the basis that rivals were not entirely excluded from the market and some rival products were priced higher than some of its products.

"The DOJ also appealed on the question of whether a finding that a firm had not violated Section 3 of the Clayton Act implied as a matter of law that the firm had not violated Section 2 of the Sherman Act. Dentsply argued that, because Section 3 is a more restrictive standard than Section 2, the Government could not, as a matter of law, prevail in the Section 2 claim having lost the Section 3 claim. The appellate court ruled that the Section 2 appeal could proceed and stand alone from the Section 3 claim that was not appealed.

The second was the claim that Dentsply was a monopolist that had successfully denied efficient distribution to rival manufacturers with clear predatory intent and no legitimate business justification.\(^5\)

In evaluating whether Dentsply had violated Section 2, the Third Circuit applied a standard that recognized the effects of weakening competition even if it were not eliminated:

Under [Section 2] of the Sherman Act, it is not necessary that all competition be removed from the market. The test is not total foreclosure, but whether the challenged practices bar a substantial number of rivals or severely restrict the market's ambit.\(^5\)

Although the product and geographic scopes of the market (artificial teeth sold in the United States) were not subject to serious dispute, the set of customers in the market was. In oral argument during the appeal, Dentsply asserted that the only sales in the relevant market were sales directly from manufacturers to dental laboratories and that sales by Dentsply to labs through dealers were not in the relevant market.\(^6\) The appellate court rejected this argument and held "that the relevant market here is the sale of artificial teeth in the United States both to laboratories and to the dental dealers."\(^6\)

It is difficult to see any economic logic in the defendant's interpretation of the relevant market. Acceptance of this argument would imply that Dentsply was not a participant in the relevant market because it made no direct sales to labs. Using this approach more broadly would imply that a manufacturer could sign long-term exclusive contracts with 100 percent of the available dealers, account (indirectly) for 99.9 percent of all sales to the ultimate consumers, and still be found to have a minimal presence and no competitive effect in "the market for direct sales."\(^6\)

For these reasons, the appellate court was correct to reject this argument. However, the court also asserted that,

[in concluding that Dentsply lacked the power to exclude competitors from the laboratories, "the ultimate consumers," the District Court overlooked the point that the relevant market was the "sale" of artificial teeth to both dealers and laboratories. Although some sales were made by manufacturers to the laboratories, overwhelming numbers were made to dealers.


Thus, the Court’s scrutiny should have been applied not to the “ultimate consumers” who used the teeth, but to the “customers” who purchased the teeth, the relevant category which included dealers as well as laboratories.

This mis-focus led the District Court into clear error.”

In other words, the appellate court believed that dealers should be considered customers and that Dentsply’s policies thus excluded rivals from serving important customers.

From an economic perspective, the appellate court’s focus was itself misplaced. The question for efficiency and consumer welfare is whether the practice at issue impeded the ability of manufacturers to sell to their ultimate customers (which, here, could be considered to be dental laboratories as surrogates for patients).’ To see why, suppose that there are two methods of distribution, A and B. If A and B are perfect substitutes, both of which are otherwise competitively supplied at constant cost, then foreclosing access to B is irrelevant to consumer and manufacturer welfare. If methods A and B are imperfect substitutes, then the welfare effects of foreclosing access to B could be understood in terms of impeding a manufacturer’s ability to sell to the ultimate consumers.

In finding that Dentsply possessed the power to exclude rivals and thus had monopoly power, the appellate court disagreed with the district court’s conclusion that rivals did poorly because of their own decisions. The appellate court stated that

‘[i]t has not been so much the competitors' less than enthusiastic efforts at competition that produced paltry results, as it is the blocking of access to the key dealers.”

The appellate court agreed with the DOJ that the at-will nature of agreements did not imply that they lacked anticompetitive effect. Rather than viewing the short terms of the contracts as mooring them in a safe harbor, the appellate court focused on "the nature of the relevant market and the established effectiveness of the restraint."' The appellate court stated that

‘[a]lthough its rivals could theoretically convince a dealer to buy their products and drop Dentsply's line, that has not occurred. In United States v. Visa USA, 344 F.3d at 229, 240 (2d Cir. 2003), the Court of Appeals held


*It is conceivable that the appellate court took dealers to be the consumers whose welfare the antitrust laws protect. That view, however, is a poor fit with the DOJ’s case, which spent little or no time considering dealer welfare except to the extent that it shed light on other issues.


that similar evidence indicated that defendants had excluded their rivals
from the marketplace and thus demonstrated monopoly power.”

As discussed earlier, the appellate court’s conclusion that the at-will nature
of Dentsply’s exclusive dealing agreements was not dispositive was in
accord with economic theory.

Lastly, the appellate court found that dealers provided certain benefits
that manufacturers could not duplicate on their own, including one-stop
shopping and aggregating tooth returns.

Based on the analysis above, the appellate court reversed the district
court to find that the exclusive dealing requirement violated Section 2 of the
Sherman Act.

THE REMEDY

The Final Judgment was issued on April 26, 2006. The Final Judgment
ordered Dentsply to eliminate Dealer Criterion 6 and further ordered that

Dentsply shall not condition the sale of its teeth or any other product to a
dealer, or authorize a person as a dealer of Dentsply teeth, based on that
dealer’s sale of non-Dentsply teeth or that dealer’s consideration to sell
non-Dentsply teeth.”

To ensure that Dentsply did not re-create the effects of exclusive dealing
agreements through other means, the Final Judgment prohibited several
potentially retaliatory practices (e.g., withholding training from dealers that
sold competing tooth lines) that might be used as substitutes for an agree­
ment. The Final Judgment also proscribed Dentsply from offering dealers
“market-share discounts,” which might otherwise induce dealers to become
exclusive distributors of Trubyte teeth.

ROLE OF INTENT

At trial, the Government introduced several exhibits indicating that
Dentsply executives intended its exclusive dealing policy to serve as a
means of weakening rivals’ ability to compete, and the district court con­
cluded that this was Dentsply’s intention. This finding raises the question of
what role intent should play in antitrust litigation.


Economists frequently assert that effects—rather than intentions—are what matter for welfare and that thus intentions are irrelevant. However, there is an important sense in which intentions can be taken as analytical findings. If one expects businesspeople to know what they are doing, then it is telling that Dentsply’s executives did not agree with Judge Robinson’s views on the viability of direct distribution or non-Trubyte dealers as substitutes for the members of Dentsply’s dealer network. Dentsply’s executives clearly had concluded that access to dealers had important effects on rivals’ abilities to compete. As the DOJ stated in its appeal,

Under the [district] court’s view, Dentsply for fifteen years has prohibited independent dealers from selling almost any competitive teeth—with the intent and expected effect of keeping out rival manufacturers—when such conduct was obviously unnecessary and unworkable, because its rivals were inept and could do just as well by selling directly to labs or using other dealers.”

Of course, as to ultimate effect, it is theoretically possible that Dentsply was doing the right thing for the wrong reason. That is, its executives may have thought they were engaging in exclusion but inadvertently preventing free-riding.” And there may have been issues relevant for antitrust analysis that either did not rise to the executives’ attention or that the executives were unqualified to analyze. Hence, intentions alone are not enough to establish anticompetitive effect. But neither are they entirely uninformative.

CONCLUSION

It is difficult for an outsider to assess a highly fact-intensive, rule-of-reason decision. One is on firmer ground assessing the economic logic of the broad principles applied. The appellate decision applied two. First, it applied the standard that exclusionary behavior that significantly raises rivals’ costs can harm competition to the point of constituting a Section 2 violation without totally foreclosing competitors.” Second, it found that even at-will agreements can have exclusionary effects.

   69 As discussed in footnote 31 supra, there are circumstances under which exclusionary practices that reduce the sales of inefficient competitors can increase efficiency. Such an argument is unlikely to be raised as an antitrust defense, however, and Dentsply does not appear to have made this claim.
60 In this regard, the case was part of an evolving standard also developed and embodied in several recent cases, including United States v. Microsoft Corp., 253 F.3d 34 (D.C. Cir. 2001) (en banc) and United States v. Visa U.S.A., Inc., 344 F.3d 229 (2d Cir. 2003), cert. denied, 125 S. Ct. 45 (2004). These cases are addressed in this volume in Case 20 by Daniel Rubinfeld and in Case 19 by Robert Pindyck, respectively. See also Jacobson (2001).
Both positions clearly are in accord with sound economics. Rivals do not have to be driven from the market in order for a practice significantly to harm competition and consumers, and even at-will distribution agreements can have the effect of inefficiently raising rivals’ costs in some circumstances. Nevertheless, there is a question as to whether the court is setting a standard that is difficult to apply in practice, and there remain complex issues regarding whether considerations of error costs might lead courts optimally to require total exclusion or the use of long-term contracts before finding an offense. Only time—and future litigation—will tell.

REFERENCES


INTRODUCTION

The economic theory of bundling has moved from the classroom and academic journals to the public policy arena. Its debut was dramatic. On July 3, 2001, the European Commission (EC) blocked the $42 billion merger between General Electric (GE) and Honeywell.1 A primary reason for its objection to this combination was a concern over bundling.

This case study uses the context of the proposed GE-Honeywell merger to address the concerns raised by bundling. We set out the theory as put forth by the EC and try to reconcile this theory with both the economic theory of bundling and the facts of the case. We discuss what is meant by “bundling” and explain when it is a potential problem and when it is not. Based on this understanding, we propose antitrust policy recommendations to deal with the novel issues raised by bundling.

BACKGROUND

On October 19, 2000, United Technologies Corporation (UTC) reported that it was in merger discussions with Honeywell. Three days later, a merger was announced—but the buyer was GE, not UTC.
This case generated a good deal of attention. General Electric is one of the most well-known and admired companies in the world. At $42 billion, this was a large merger even for GE. The proposed integration passed the scrutiny of the U.S. Department of Justice (DOJ). Because of the size of GE and Honeywell's European sales, the merger also had to be approved by the European Commission. On July 3, 2001, that permission was denied. The divergence of outcomes between the U.S. and European antitrust authorities added to the publicity of this case.

THE PLAYERS

GE's 2001 revenues exceeded $125 billion, and its businesses included everything from plastics and television (NBC) to financial services, power systems, medical imaging, and lighting. In the arena of aviation, GE produces aircraft engines on its own (GEAE) and through CFMI, a 50-50 joint venture with SNECMA (a French company). This joint venture accounted for a large majority of GE's engine sales, as CFMI is the exclusive provider of engines for Boeing's most popular plane, the 737. CFMI engines also power the Airbus A320 family and the A340-200/300. GE's own engines power the Boeing 777, 767, and 747 planes. They also power the Airbus A300, A310, A330, and the not-yet launched A380 super-jumbo aircraft.

Honeywell started out in heating and environmental controls and over time developed a leadership position in aerospace. Honeywell's position in avionics was enhanced through a series of mergers, most notably the purchase of Sperry Aerospace in 1986 and a merger with Allied Signal in 1999. Allied Signal was itself a leader in aerospace, the result of Allied Corporation's purchase of Bendix in 1983 and merger with the Signal Companies in 1985. Along with avionics, Honeywell's nonaircraft aerospace products include auxiliary power units (which give power to the

---

2The Commission has authority to review all mergers, acquisitions and takeover bids and other deals that can be defined as a 'concentration', involving companies with a combined turnover worldwide in excess of €5.0 billion and European sales of at least €250 million' (Commission press release IP/02/393). See also Article 1 (2) (a) and (b) of the European Merger Control Regulation, Council Regulation 4064/89 EEC of 21 December 1989 on the Control of the Concentration between Undertakings, 1990 O.J. (L 257) 13, as amended by Council Regulation 1310/97 EC of June 30, 1997, 1998 O.J. (L 180) 1.

3The parties appealed the decision to the Court of First Instance of the European Communities. That court's decision in December 2005—adverse to GE-Honeywell—is discussed at the end of this chapter.

*The GP7000 engine, designed to power the Airbus A380, is a joint venture between GE and Pratt & Whitney.

*Honeywell instruments helped guide Apollo astronauts Neil Armstrong and Buzz Aldrin's landing on the moon.
plane when on the ground), starter motors, environmental control systems, aircraft lighting systems, engine accessories and controls, wheels, and braking equipment. In 2001, nearly one-half of Honeywell's $23 billion of revenue came from its aerospace division.

THE CASE AGAINST THE MERGER

To block a merger, the Merger Control Regulation requires the Commission to demonstrate that the proposed merger would lead to market dominance. According to European case law, dominance is defined as:

a position of economic strength enjoyed by an undertaking which enables it to prevent effective competition being maintained on the relevant market by giving it the power to behave to an appreciable extent independently of its competitors, customers and ultimately of consumers.

In this case, the claimed route to market dominance was unusual. It was not through the merger of competitors. Nor was it through a vertical integration of customer and supplier.\(^6\) Given the depth and breadth of GE and Honeywell product offerings, there was a remarkable lack of overlap between the two companies. Instead, the focus of the merger review was on "conglomerate effects" or horizontal integration issues of bringing complements together.\(^7\)

The Commission's case against the merger emphasized three linked points.

- First, the Commission claimed that GE had a dominant position in aircraft engines for large commercial aircraft and Honeywell had a leading position in avionics and nonavionics areas.
- Second, the Commission claimed that the proposed merger would allow the new firm to bundle these complementary products. The bundling strategy would lead to price discounts that would give the firm an unbeatable advantage over its rivals.


The very few areas of competitive overlap could have been resolved by divestitures. GE's role in speculative aircraft leasing and Honeywell's starter motor business did bring up some issues of vertical integration, which are not discussed here. See Pflanz and Caffarra (2002) for an excellent analysis of the unorthodox arguments made regarding the role of GE's leasing company, GECAS. A good summary of the commission's case regarding GECAS is presented in Giotakas et al. (2001).

The Commission referred to this as a conglomerate effect. I prefer the term horizontal integration of complements as it emphasizes the specific nature of the relationship—namely that their two products are used together by a common customer. Honeywell's customers are primarily the airframe manufacturers and airlines, not the engine maker. Thus, Honeywell is a complementor to GE as opposed to a supplier. (See Brandenburger and Nalebuff [1996] for a more formal definition of complementors.) Starter motors (and a few related components) are the exception to this rule.

Page 418
• The third leg of the argument was that this advantage would lead to the exit of rivals and thus ultimately to strengthening the dominance of the merged GE.

As explained in the EU Final Decision (para. 355):

Because of their lack of ability to match the bundle offer, these component suppliers will lose market shares to the benefit of the merged entity and experience an immediate damaging profit shrinkage. As a result, the merger is likely to lead to market foreclosure on the existing aircraft platforms and subsequently to the elimination of competition in these areas.

To an economist, the Commission’s case was unorthodox. The concern was that under the merger prices would fall—not that prices would rise. The idea that bundle discounts could be an anticompetitive strategy was novel to this case. A merger that created cost savings that resulted in lower prices would be permitted. But if the lower prices came from pricing efficiencies, this was viewed as anticompetitive. There was no discussion of whether the lower prices on the road to market dominance would or would not result in a net increase in the present discounted value of consumer surplus. Such considerations do not appear to be within the Commission’s mandate.

Even to a lawyer, the Commission’s case was unorthodox. The Commission did not demonstrate that the combination would lead to dominance, as required by its mandate. Rather, it emphasized the theoretical potential for future anticompetitive behavior.

In this case study, we focus on the Commission’s arguments concerning bundling. Not only is this aspect of the case novel, it also has wide-ranging implications for other mergers. We present the Commission’s case for why bundling might create an antitrust issue, and then discuss whether the theory and evidence justified its conclusion.

It will help to define bundling at the outset. Many items are sold together in a package. That in itself does not mean that the items are bundled. The answer depends on how else and at what prices the items can be purchased. For example, if two items are only sold together and are not available separately, this is a case of pure bundling. This may be done via pricing or via technology.

‘Range effects had been considered in a number of prior EU cases, including Coca-Cola/Amalgamated Beverages and Guinness/Grand Metropolitan. These cases focused on distribution efficiencies rather than pricing strategies.

“Schmitz (2002) criticizes the decision on this basis: ‘Although it is probably true that the new company would indeed have the potential to bundle and it cannot be ruled out that at one point in time it might engage in this behavior, using this potential to conclude that the merger would strengthen a pre-existing dominant position within the meaning of Article 2 European Merger Control Regulation is questionable. . . . Describing the question of whether it is permissible to block a merger because of possible future bundling as theoretical, hardly fits the impact it has. . . . [T]he tool for this investigation is and must be Article 82 EC, not the European Merger Control Regulation.’

Page 419
Pure bundling becomes tying when one of the items, say, item 1, is sold by itself, but item 2 is available only as part of a package with 1. Thus, the National Football League (NFL) has a season pass program to television viewers to watch every game of the season. This viewing option is available only on satellite TV and is not offered to cable viewers. Of course, satellite viewers can also buy a contract without the NFL season pass. Thus, we would say that the NFL season pass is tied to satellite TV.

The most general form of bundling is mixed bundling, and this was the type of bundling emphasized by the Commission. In mixed bundling, two items are available both separately and as a 1-2 package. What makes this bundling is that the package is sold at a discount relative to the individual items. If the package is simply priced at the sum of its component prices, and these components are each available on an à la carte basis, then we do not call this bundling, as there is no strategic impact of the bundle pricing. Note that mixed bundling includes pure bundling and tying as special cases.

In this next section, we first present the argument as made by the Commission concerning mixed bundling and then question whether this theory of bundling was applicable to the case at hand.

THE ECONOMIC THEORY OF BUNDLING

The theory of bundling begins with Cournot (1838). Cournot considered the case of a monopoly seller of good 1 and a different monopoly seller of good 2 where the two goods are used together by the customer. Cournot used as an example the case of copper and zinc that are combined to make brass. Here one can think of the two goods as a jet engine and avionics.

Cournot’s insight is that two monopolists, acting independently, will set an inefficiently high price. Were they to merge or coordinate their pricing, they would lower their prices and earn more money. The simple intuition is that the lower price of good 1 stimulates sales of good 2 (and vice versa) and this effect is not considered when goods 1 and 2 are sold independently.

It is not surprising that the merging firms make more money. What is unusual here is that prices fall so that consumers are also better off. The merger is a Pareto improvement. Thus, the antitrust authorities should encourage such mergers. We return to this issue when we discuss policy implications.

The Cournot example is the horizontal equivalent of “double marginalization.” Each firm causes a negative externality on the complementary

"One of the ironies of this case is that if one took the view that GE and Honeywell each had a monopoly position, then bundling would unambiguously improve welfare. The only possible source of harm would be on competitors. But if each firm (as a monopolist) does not face competitors, there is no harm done.

Page 420
products by raising its price. When the two firms combine, they internalize this effect and lower prices. The simplicity of Cournot’s argument has led to general confusion about when the theory is applicable.

In some ways this result is very general. It does not depend on the specific form of the demand function nor the cost function. It does not require that the goods be perfect complements. But there are several hidden assumptions on which the result relies. Two, in particular, are of concern to us.

First, the basic Cournot model does not consider the impact of the merger on any other firms in the market. This is by design. In Cournot, the producers of goods 1 and 2 are alone in the market. To apply this approach to the facts as interpreted by the Commission in the GE-Honeywell merger, we need to consider how the results change when the two merging firms are not alone in the market. Naturally, the results are more complicated.

There are now two reasons to cut price: market expansion and competition with rivals. Cournot looked at the price reduction as a way to increase the total market. Though the total demand for airplane engines can be expanded with a price reduction, the engine (along with avionics) is only a minority of the total airplane cost. The greater impact of a price cut is through the potential to gain market share from rivals.

Because there are rival firms, there will also be a response to a price cut. This response may offset the potential gain to the merging firms. Thus, we will want to consider the equilibrium impact on the nonmerging firms and on consumers to determine the overall social welfare implications.

A second reason why the Cournot framework may not apply to the GE-Honeywell merger is that the basic result depends on an unstated assumption: that firms set a single price in the market to all customers. This is a quite reasonable assumption for a typical consumer good, such as Microsoft Office. But it is not a reasonable assumption for the sale of large commercial products in which the two parties engage in extensive negotiation as part of the sale process. If firms can price discriminate or negotiate with each customer, then the advantage to bundling disappears.

Our focus here will be on the applicability of the Cournot model. There are other reasons to offer a bundle besides curing double marginalization. For example, even a monopolist seller of both good 1 and good 2 might be able to use bundled pricing as a way to improve its ability to engage in price discrimination." This incentive to bundle is more important if there is demand for the goods individually (as is the true case with copper and zinc) than when all customers buy both goods (as is typically the case with jet engines and avionics)." We do not emphasize this aspect of bundling, as it did not play a role in the GE-Honeywell case.

"See, for example, Adams and Yellen (1976); McAfee, McMillan, and Whinston (1989); Bakos and Brynjolfsson (1999).

"There is still an advantage to bundling in a world where all customers buy a 1-2 package and the rival sellers of good 1 (and good 2) are imperfect substitutes. This advantage exists only so long as the firm cannot perfectly price discriminate."
COMPETING AGAINST BUNDLES

We first extend the original Cournot model to cover the case where the two sellers face competition in the market. The results from this approach were at the heart of the Commission’s theoretical argument against the merger.

This approach is based on Nalebuff (2000). We consider the case with four firms in the market. There are two differentiated versions of the first good, produced by firms \( A_1 \) and \( B_1 \) and two differentiated versions of the second good, produced by firms \( A_2 \) and \( B_2 \).

To model the product differentiation we assume that both \( A \) components are located at 0, while both \( B \) components are located at 1. (These locations “0” and “1” can be thought of as points in geographic space or in product characteristics space.) The two goods can be thought of as avionics and engines with the “\( A \)” firms being GE and Honeywell and the “\( B \)” firms being Pratt & Whitney and Rockwell Collins.

Because planes require both engines and avionics, the customer will purchase one unit of both components in the basket. Thus, each customer will buy one of (\( A_1, A_2 \)), (\( A_1, B_2 \)), (\( B_1, A_2 \)), or (\( B_1, B_2 \)).

Customers assemble the package that best suits their preferences. Each customer purchases the package with the smallest total cost, where total cost is composed of price plus a linear unit transportation cost. We assume that for each of the two goods, the customer’s ideal location is uniformly distributed over \([0, 1]\). Location preferences for each good are independent.

There are three possible market structures. All four firms act independently; the two \( A \) firms bundle as do the two \( B \) firms, resulting in bundle-against-bundle competition; the two \( A \) firms combine, while the two \( B \) firms remain separate.

Case 1: \( A_1, A_2, B_1, B_2 \), each separate
Case 2: \( A_1-A_2 \) combination versus \( B_1-B_2 \) combination
Case 3: \( A_1-A_2 \) combination versus \( B_1 \) and \( B_2 \) separate

Case 1: All Firms Act Independently

With a uniform distribution of customers and unit transportation costs, the equilibrium prices equal 1, and the market is evenly split between firms \( A \) and \( B \).

Later we will explore cases where the sellers have more information as to the customer’s preferences. This leads to a different distribution of customer locations.

The case where the two \( B \) firms combine and the \( A \) firms remain separate leads to the same results.

This result assumes “Bertrand” price competition between each of the pairs of competitors. We further assume that production has constant and equal marginal costs. With this assumption, profit margins are independent of costs, and so we employ zero marginal costs in the results below.
where $P_i$ is price and $\pi_i$ is profit.

This case is the baseline from which we can evaluate the impact of coordinated pricing decisions. In this baseline case, consumers mix and match their preferred components and pay a price of 2 for their two-good customized bundle.

**Case 2; Bundle versus Bundle**

Here all the A firms coordinate their pricing and sell their product as a bundle against the B firms, who have also coordinated their pricing decisions. Let bundle A sell for $P_A = P_{A1} + P_{A2}$, and bundle B sell for $P_B$, defined similarly. In this equilibrium

$$P_A = P_B = 1; \quad \pi_A = \pi_B = \frac{1}{2}.$$ 

Profits fall by 50 percent. This is because the aggregate bundle price has fallen by 50 percent. The price of the entire bundle is reduced to the prior price of each of the single components. In hindsight, the intuition is relatively straightforward. Cutting price brings the same number of incremental customers as when selling individual components. So the bundle price must equal the individual price in a symmetric equilibrium. Bundle against bundle is ferocious competition. Similar results hold for bundles with more than two goods as demonstrated in Nalebuff (2000).

**Case 3: Bundle against Components**

The pricing externality suggests that the bundler will have an advantage over the component sellers. But the results of case 2 suggest that this gain may be offset by an increase in competition induced by the A firms selling their products only as bundle. Which effect dominates?

It turns out that the increased competition effect dominates so that bundling reduces Firm A’s profits from 1 to 0.91. The bundler does roughly 50 percent better than the sum of the uncoordinated B firms. The market share moves from a 50-50 split to 63-37. But even with this gain in share, the bundler does about 10 percent worse than in the case where each component is sold in an uncoordinated fashion. The explanation is that the bundle takes away enough market share from the B firms so that the resulting

\[ P_A = 1.45; \quad P_B = 0.86, P_B = 1.72; \quad \pi_A = 0.91, \pi_B = 0.32 \]

*For simplicity, we assume that consumers buy only one of the two bundles.*
equilibrium prices are low enough to make the A firm worse off. Thus, even though it leads to an advantage, there is no incentive to bundle.

These results, taken from Nalebuff (2000), consider the case of pure bundling—products A and A are available only as a bundle. To apply this model to GE-Honeywell, the results need to be extended to cover the case of mixed bundling.

Mixed Bundling

Rolls Royce presented to the Commission an extension of this model to include the case of mixed bundling. While it is difficult to find a closed-form solution, the approximate equilibrium prices and profits can be found through simulation. The results are not identical to the pure bundling case, though they have the same flavor.

The A bundle is sold for a 19 percent discount below the premerger price of 2, and the component prices for A rise from 1 to 1.21. The B firms respond to this increased competition by lowering their component prices to 0.89.

\[ P_A = 1.63, P_{A1} = 1.21; P_{B1} = 0.89, P_B = 1.78; \pi_A = 0.97, \pi_B = 0.40 \]

Even with mixed bundling, the merging firm still sacrifices profits. Profits fall from 1.00 to 0.97 or 3 percent. Although its profits fall, firm A gains an advantage over its rivals. Firm A’s market share is 55.4 percent, and the rival’s profits fall by 21 percent.

This model is obviously quite stylized. The results about whether or not the bundling is profitable can be reversed with relatively minor changes in parameter values or modeling assumptions. For example, bundling becomes more profitable the more items are added to the bundle (at least for the case of pure bundling).

While the biggest impact from lowering price comes from gaining market share, there is also the potential to expand the total market. As recognized by Choi (2001), even if this effect is small, it can be enough to make bundling profitable.

Thus, the Commission reached the conclusion that economic incentives would lead a firm to engage in mixed bundling. That a firm that bundles obtains an advantage over its rivals is a relatively robust conclusion. But whether or not a multiproduct firm has an economic incentive to bundle is a much more delicate finding. And by the same token, so is the expected loss to the competition. Among other factors, it depends on the

---

"As the bundle grows in size, the gap in prices continues to grow and, consequently, so does the bundler’s market share. Once the bundle has four or more items, equilibrium profits rise for the bundling firm. See Nalebuff (2000).

"The nonconfidential version is presented in Choi (2001).
number of items in the bundle and the elasticity of total market demand. The results presented rely on a specific distribution of preferences, namely a uniform distribution. It is not clear if the results would be robust across different distributions of customer preferences.

The results also depend on the two goods in the bundle being of symmetric importance to the consumer. Clearly, the engine is more important than the avionics or even all of the potential Honeywell components combined.

An aircraft engine might have a $15 million price tag, while a piece of avionics could sell in the $100,000 range. Overall, the avionics components add up to less than 5 percent of the total aircraft cost. The basic model of bundling has the two products being symmetric in valuation. Nalebuff and Lu (2001) extend the earlier model to allow for asymmetry in importance. In the examples considered, there appears to be little incentive to bundle and minimal impact on competitors. In fact, with pure bundling and enough asymmetry, bundling actually increases the profits of all the players in the market.

The Rolls Royce model examines only the two-goods, two-vendor case. Realistically, the number of avionics and nonavionics goods purchased is in the several dozens, and the number of important vendors is at least a dozen. We need to understand not just whether mixed bundling is attractive or not, but also how much more attractive it is as the bundle size increases and what impact that increased bundle has on the market.

Thus, before we try to use these types of models to make predictions about the likely impact of a merger, the models need to be robust enough to capture some important elements of the real-world market. If one takes the results of this model at face value, it is not clear why there should be any antitrust concern. Average prices fall in the market. Consumers are better off at the expense of firms.

One can make the argument that social welfare falls, but as Patrick Rey observed in his presentation to the European Union Merger Task Force this is truly an artifact of the model. As demand is inelastic, the only change in social welfare is due to a change in transportation costs. Since transportation costs are minimized in the initial symmetric equilibrium, any change would lead to a fall in social welfare. But this argument depends critically on the starting points being completely symmetric. If, for example, the merging firms have a superior position and a larger share, then the bundle discount can lead to increased efficiency.

"But even assuming for a moment that all of these issues did not exist, work by Patrick Rey shows that even on its own terms the Rolls Royce model was fatally flawed. In order to estimate the impact of the proposed merger, one needs to have coefficients on the model. The problem is that there is only one observation in use to estimate these parameters. Rolls Royce used the existing market shares. At best that would leave the model exactly identified. However, in this case the model was still underidentified. Thus, the impact of the merger would depend on the choice of some parameters for which there were no data to make an estimate."
The Commission’s problem with bundling was not with the immediate loss to social welfare but rather with the long-run (and even short-run) impact on competition. Similar to the argument against predation, the Commission believed that rivals would exit and that GE-Honeywell would obtain and exploit a dominant position.

**Dynamics**

If one wants to consider the expected market dynamics, then there are other factors that should be taken into account. For example, is it realistic to imagine that the individual competitors would not respond in any way? One option for those firms is to invest in product improvements or cost reductions. Another option would be to offer a competing bundle.

For the B firms, offering a competing bundle would lead to a further reduction in profits, but it would also level the playing field (recall case 2). Firms may prefer to be in a symmetric position relative to a rival. They may prefer a lower but level playing field if they are worried that a bundler would use its profit advantage to position itself better in R&D or gain other advantages in a repeated game.

Even if the B firms did not want to offer a competing bundle, their customers could drive them to this outcome. Customers would stand to gain a great deal if they could create a bundle-against-bundle competition. Customers are anything but passive in this market and would use their power to influence the nature of competition.

The Rolls Royce model looked at the advantages of bundling when the competition was selling their products individually. Whatever advantages may exist, they quickly disappear if the rival firms coordinate and offer a competing bundle. Given the level playing field for rivals and the advantages to customers, a firm that introduces a product bundle cannot expect that rival firms will not offer competing bundles.

**NEGOTIATING BUNDLES**

The previous discussion suggests several potentially interesting questions for antitrust policy, but its relevance to the GE-Honeywell merger is doubtful. The reason is that the previous results on bundling all depend critically on the assumption that there is one price to all customers in the market. This is such a basic assumption to most economics models that it is usually not even stated. However, this assumption does not apply to the aerospace industry.

Customers don’t pay list price for jet engines or avionics. Airplane customers are large and powerful. A vendor cannot ignore an airline that asks for a better price. Nor are vendors uninformed as to their customers’
preferences. The vendors in this market spend a great deal of resources getting to know their customers. Vendors take into account previous purchases as well as technical performance differences between their products and those of competitors. Going into a competition a vendor has a good idea of where it stands; and by the end of the competition, it has a very good idea.

The result of this buyer power and vendor information is that prices are negotiated, not set by the seller. Examples of transaction prices confirm the fact that different buyers pay different prices. Every deal is negotiated, and the price is customized to the specifics of the situation.

In a world where firms negotiate prices with customers and do so with perfect information, the combination of two complementor firms is completely neutral. To see why, first take the case where a customer has a preference for the two firm A products. Imagine that the net advantages are equal to (0.2, 0.3). In this case, the A firms would be expected to win both competitions, whether the items are sold separately or if the firms merge and the products are sold in a bundle.

Before conceding defeat, the B firms would be willing to price down to marginal cost. In the case where the A goods are sold separately, firm A would be able to negotiate a profit up to 0.2 on good 1 and firm A could negotiate a profit of up to 0.3 on good 2. If the goods were sold as a bundle, then the merged firm A would be able to negotiate a profit of up to 0.5 on the bundle.

Nothing is different if firm B has an advantage in one (or both) of the goods. Take the case where the customer has a net preference for A products of (—0.2,0.3), so that the customer actually prefers B to A. With individual pricing, firm A would lose to B, and firm A, would beat B. If merged A firm tried to sell the two products as a bundle, it could do so, but only at a profit of 0.1. This would be worse than selling just the second good at a profit of 0.3.

These two examples are quite general. The point they illustrate is the following: When the customer type is known and prices are negotiated, bundling can never lead to higher profits. If the customers would have made all of their purchases from a single firm, bundling has no impact—on customers, prices, profits, or efficiency. If customers would prefer some products from A and others from B, then the combined firm will continue to offer the individual goods at their premerged prices. Forcing a bundle on the consumer can only lower firm A’s profits. In effect, it would have to subsidize its disadvantage using profits it could have earned from products where it is strongest. This is no different from selling individual components at a loss—a strategy it can do but would choose not to, even without bundling.

One additional perspective can help with the intuition. Firms make profits only to the extent that their products are differentiated. Profits exist to the extent that the firm has an advantage with the customer. When a firm bundles two good products or two bad products together, the advantages
(or disadvantages) sum up and there is no impact. But when a firm mixes good and bad products together, this mitigates the advantage and profits fall accordingly. With mixed bundling, there would be no bundle discount and thus no effect at all.

The perfect information negotiation model is designed to capture the basic nature of competition in this market. But, like all models, it presents a simplified description of the market. While vendors are well informed about the customer, their information is not always accurate. Firms can still negotiate prices even with good but imperfect information as to customer preferences.

The conclusion from this negotiation model is not a narrow result. The mathematics of the more realistic cases becomes more difficult, but simulation results suggest that bundling has little effect when vendors have good but not perfect information.

As one example, imagine that the firms don't know the customer's exact location (and thus preferences) but do know which firms the customer prefers. Consider a customer who is known to prefer both A products.

In this case, bundling offers a small advantage. A firm that knows it has an advantage in all components can use a bundle to do a better job of price discrimination. Market efficiency increases. Competitors' profits fall, but their profits in this case were very low to start with.

In contrast, bundling is not profitable when the merging firm is at a disadvantage in all products. Here we also note that this is the case where rivals have the highest profits. Thus, bundling has the lowest impact when rivals make the most money.

Pure bundling or tying would be counterproductive in the two cases where the merging firm is better in one component and worse in the other. If the merging firm employs mixed bundling, then the majority of consumers do not take advantage of the bundle, and there is again a small impact of the mixed bundle.

Averaging across these four cases, the net impact of the mixed bundling strategy is reduced by 50 percent. This is in line with the improved information. Half of the uncertainty has been removed in the sense that each firm knows if it is ahead or behind, although not by how much. Nalebuff and Lu (2001) show that with even better information, the impact of mixed bundling is even smaller.

The distribution used in generating the simulation results provides the results for a given type of customer with some uncertainty, but that leaves open the question of what is the proper distribution of customer types. The tightness of the distribution is a proxy for the quality of information in any particular negotiation.

21 For more specifics on this approach, see Nalebuff and Lu (2001).

22 Another way of putting this is that each firm knows which half of the line the customer is in and thus the range of uncertainty has been reduced by half.
The wide margin variations that we observe are indicative of high-quality information, but this is not something that has been empirically measured and calibrated to the model.

EMPIRICAL EVIDENCE OF DOMINANCE AND BUNDLING

Ultimately, the Commission’s case against the proposed merger rested on a claim of market dominance. The starting point in this argument was that GE already had a dominant position in aircraft engines for large commercial aircraft. Recall that market dominance means that a firm can act independently of its rivals and its customers.

It almost follows from the definition that a firm without a commanding market share cannot have a dominant position in the market. The Commission presented its calculation that GE had a dominant position with a 52.5 percent market share of the installed base of engines on large commercial aircraft still in production:

<table>
<thead>
<tr>
<th>GE</th>
<th>PVMAE</th>
<th>RR/IAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrow Body</td>
<td>51</td>
<td>22</td>
</tr>
<tr>
<td>Wide Body</td>
<td>54</td>
<td>31</td>
</tr>
<tr>
<td>Overall</td>
<td>52.5</td>
<td>26.5</td>
</tr>
</tbody>
</table>

Planes still in production leaves out planes still in service but no longer produced. While new engines are obviously no longer sold to planes no longer being built, this perspective misses the spare parts market. With spare parts an engine can be “sold” up to ten times over its working life. Pratt & Whitney, in particular, has a large annuity coming to it from selling spare parts to planes in service that are no longer being produced. Using the planes-in-service approach, GE’s market share falls to 41 percent.

The calculation of market shares for PW (Pratt & Whitney) and RR (Rolls Royce) include the engines of IAE, a joint venture between Pratt & Whitney and Rolls Royce. IAE’s market share is split evenly between PW and RR. In contrast, all of the market share of CFMI, the 50-50 joint venture between GE and SNECMA, is attributed to GE. If CFMI is eliminated from consideration and we consider market share on planes in service, then GE’s market share falls to 10 percent. If we attribute half of

“The data are of December 31, 2000. The table is reproduced from European Commission decision in Case No. COMP/M.2220—General Electric/Honeywell, para. 70.

“The Directorate-General Competition justified this calculation: “Although in legal terms GE and SNECMA jointly control CFMI, the only meaningful attribution of market shares for the purposes of analyzing the transaction could only be made to GE, to the extent that SNECMA is not an independent supplier of civil jet engines for large commercial aircraft. The analyses of the joint venture and of SNECMA’s participation in other GE engine programmes indicated that SNECMA would act jointly with GE as a profit-maximising entity” (Giotakos et al. [2001]).
CFMI to GE and half to SNECMA, then GE’s market share is still only 28 percent. Even if we use the planes in production and credit only half of CFMI to GE, then GE’s market share is 36 percent.

In any of these calculations other than that of the Commission, GE’s limited market share would practically preclude it from having a dominant market position. Which definition is correct?

There is no one correct definition of market share made in the abstract. The issue should be market share measured for what purpose. For example, if the purpose is to evaluate the firm’s financial resources to continue investing in this business, then all of its revenue streams are relevant. Thus, PW can use earnings from its entire installed base (not just planes in production) to finance new investment. GE only gets one-half of the revenue from CFMI. Thus, for this purpose it would seem that the 28 percent number is most appropriate as it reflects GE’s revenue share of today’s market.

To understand future financial viability, one might also want to examine preorderds on next-generation planes, those that have been launched but are not yet in production (such as the A380, B777x). It is common for airplanes to be ordered in advance of production. As of 2001, Rolls Royce had a 40 percent share of the engines on these next-generation planes, while GE and CFMI’s combined share was second with 38 percent, and PW was third with a 21 percent share.

If market shares are used to provide evidence of market power, it would seem that most of CFMI’s share should not be included in GE’s share. This is because the great majority of CFMI’s sales are not in a position to exercise any market power. Recall that CFMI is the exclusive provider of engines to the Boeing 737 plane. Prior to entering this exclusive relationship, Boeing realized that its customers would be in an untenable bargaining position if CFMI were the sole provider. Thus, Boeing negotiated a deal with CFMI as part of the exclusivity contract.

For most engines, the airline purchases the engines separately from the plane. Having decided upon a 747, the airline can then put the engine order out to bid between GE, PW, and RR. But in the case of a 737 purchase, the customer negotiates with Boeing and the engine is included in the price. CFMI does not have the ability to control engine pricing on these orders. For that reason, a more appropriate measure of GE’s market share used to measure market power would be its share of engines excluding its exclusive-contract sales. This corrected share is on the order of 10 percent or 20 percent, depending on whether one looks at planes in service or just those in production.

“In addition to CFMI’s exclusive sales contract for the Boeing 737, its CFM56 engines power the Airbus A320 family and A340 long-haul.

“There are some GE exclusive sales, such as for the two new ‘longer range’ (LR) versions of the Boeing 777. Here, again, the engine is sold at prenegotiated terms. One would want to take out all of the exclusive engine sales in calculating a firm’s ability to act strategically.}
If market shares are used to understand the potential for bundling, then, again, it would seem that CFMI’s sales should not be attributed to GE. This follows for two reasons.

First, because the engine is offered to Boeing at a predetermined deal it is almost impossible to employ a bundled pricing strategy. If the customer has already bought the Boeing plane and CFMI engine, there is no gain in providing a retroactive engine discount for the purchase of Honeywell components. The company would do just as well to offer the discount directly on the Honeywell parts.

The only way even to offer a bundle discount would be to promise a future discount on Honeywell components if the 737 plane is purchased. This would make sense only if the discount would increase the 737 plane sales. (If the sales of 737’s are unchanged, then there is no increase in demand for the CFMI engine, and hence the discount might as well be applied directly to the Honeywell components.)

There are practical problems with this approach. The cost of Honeywell components is small relative to the plane and engine. Thus, there is little ability for the tail to wag the dog. Second, the sale of avionics and other Honeywell components is done at a future time at a negotiated price. One party can’t promise to give the other party a “better deal” in a future negotiation, as there is no baseline against which to measure what makes a better deal.

Bundling would be further complicated by the fact that CFMI is a joint venture. Thus, the decision on offering a bundle price would have to be made by both GE and its partner SNECMA. By its charter, CFMI is always led by a SNECMA representative. The joint venture nature of the relationship complicates CFMI’s ability to offer a bundle, as SNECMA has no incentive to help sell Honeywell avionics.

GE’s ability to act strategically so as to gain market share is practically restricted to its own engine sales, and this is roughly 20 percent of the market for planes in production. A 20 percent market share does not put a firm in a position where it can act independently of its rivals and customers.

However one measures shares, there is no disagreement regarding the extensive use of bidding competitions, the emphasis on market intelligence, and the very dynamic nature of shifting market shares. But the same facts were viewed very differently on the two sides of the Atlantic. In the United States, the fact that GE had won several of the recent engine competitions was viewed as evidence of competition at work. In Europe, these recent wins were viewed as evidence of dominance.

“To help make this clear, consider the following hypothetical bundling offer: If “New Air” purchases a 737 plane, GE will offer a $1 million rebate good toward the purchase of Honeywell avionics. When it comes time to negotiate the price of those avionics, Honeywell and the customer both factor in the $1 million discount when setting the price. Thus, if the negotiated price would have been $4 million, the new negotiated price would be $5 million, which achieves the $4 million post-rebate price.”

“For a more detailed comparison of the U.S. and European approaches, see Patterson and Shapiro (2001) and DOJ (2001).}
A contributing factor to GE’s alleged dominance in engines was the company’s unique financial strength. Excerpts from Giotakos (2001) demonstrate this perspective:

GE Capital offers GE business enormous financial means almost instantaneously and enables GE to take more risk in product development than any of its competitors. . . . GE has also taken advantage of the importance of financial strength in the industry through the use of heavy discounts of the initial sale of the engine. . . . Thanks to its financial strength and incumbency advantages as an engine supplier, GE can afford to provide significant support to airframe manufacturers under the form of platform-programme development assistance that competitors have not been historically in a position to replicate. . . . Unlike any other engine manufacturer, GE can afford to pay for exclusivity and capture aftermarket, leasing and financial revenues.

Patterson and Shapiro (2001) show the dangers in the approach. They look at these same activities as procompetitive. Taking risks leads to innovation; discounting benefits customers. In the United States, entrenchment of a dominant firm is a discredited theory and is no longer grounds for challenging nonhorizontal mergers (DOJ 2001):

Challenging a merger because it will create a more efficient firm through economies of scale and scope is at odds with the fundamental objectives of antitrust law. And there is no empirical support for the notion that size alone conveys any significant competitive advantage that is not efficiency-related.

Whether or not GE was starting from a dominant market position, the Commission was concerned that this merger would allow GE to extend its dominant position to Honeywell products. The proposed route to this dominant position was through bundling. To what extent was this theoretical concern supported by evidence?

While the theory suggests that bundling is unlikely to be an important factor, there remains the empirical question: Do we see much evidence of bundling in aerospace? At first glance, the answer seems to be yes: Many bids are multi-item bids. Companies such as Honeywell and its competitors often make a bid to supply a long list of components. This led several observers and even industry participants to conclude that bundling was a feature of this industry.

However, the long list of items is also broken down into prices for each individual component. And the component prices add up to the “bundle” or package price. If there is no discount, then we don’t consider this to be bundling.

Compare this situation to Microsoft Office. Microsoft’s July 2001 list price for Office XP Professional was $547. One could also buy the
components separately, but one wouldn’t. Word, Excel, PowerPoint, Access each cost $339 separately, and Outlook was a bargain at $109. The total adds up to $1,465. The software package comes at a 60 percent discount compared with the individual items.

In the cases cited by the Commission, the claimed bundle discounts were much smaller—by an order of magnitude. The claimed bundle discounts were also much smaller than those predicted by the Rolls Royce model. Even more importantly, these discounts were evidence of negotiation and not of mixed bundling.

In cases where the Commission attempted to document examples of a mixed bundle discount the analysis failed to distinguish between a discount and a discount conditional on buying a package. As a result, the cases cited actually demonstrated the absence of mixed bundling.

Even if a discount were offered for buying a package, this does not mean that the discount would not be applied to individual items. In general, this would be a matter of speculation. However, in the cases cited by the Commission, the alleged mixed bundle failed to induce the customer to buy the entire package. Thus, it is possible to observe what prices were paid and whether or not the discount was in fact conditional on buying the entire package. In the cited cases, the customer was not penalized for breaking apart the bundle. Whatever discounts the customer was offered for purchasing the bundle were applied on a proportionate basis to the partial bundle that was ultimately purchased.

The evidence of bundling provided by the Commission undercut its own argument. If bundling was such an anticompetitive tool, then why did it fail to get customers to purchase products for which the firm was otherwise at a disadvantage? The fact that customers were able to get any discount offered without having to purchase the bundle confirms the industry perspective that it is a mistake to offer a bundle discount as this will simply end up coming back as a discount on whatever the customer ends up buying.

If bundling is to be a matter of concern, we should see contracts that are won, not lost. And these contracts should offer a substantial discount for buying the entire package over a la carte purchases.

**THE (IM)PRACTICALITY OF BUNDLING ENGINES**

The theory suggests that bundling will not lead to an advantage with negotiated prices. The evidence suggests that bundle discounts are not prevalent, if they even exist. But perhaps the combination of GE and Honeywell would create a new opportunity to offer an engine and avionics/nonavionics bundle. Even if this were desirable, there are institutional features of the airplane purchase process that make this type of bundling impractical.

We have already discussed the problems of bundling CFMI engines with Honeywell’s avionics (and other nonavionic) equipment. These include
the joint venture with SNECMA and the predetermined engine price for 737s. An additional factor that makes it impractical to bundle engines with other components is the timing of purchases. Typically, the engine choice is made well in advance of the choice of other components, such as avionics.

To see why this is a problem, consider how bundling would have to work. Once GE has won the engine competition, there is no incentive to give retroactive discounts on engines if the customer would also buy Honeywell avionics. This would be no different than giving a discount on the Honeywell avionics directly.

In order for there to be any possibility of bundling to work, it must be the case that at the time the engine selection is being made the customer is led to believe that by choosing the GE engine there will be a better price on Honeywell avionics. For example, the customer might be given a 10 percent discount for its later purchase of avionics if it has a GE engine. While this story is possible in some industries, it is not applicable to avionics. The reason is that all prices are negotiated so that a discount from list price has no bite. One person can’t promise to give the other party a “better deal” in a future negotiation as there is no baseline against which to measure what makes a better deal. Here again, we see the importance of taking into account whether prices are fixed or negotiated.

The bundling of avionics components would seem to be much more practical than bundling engines and avionics. Indeed, this issue was considered only two years earlier when the Commission approved the merger of Allied Signal and Honeywell. While there was some dispute as to whether bundling exists at all, even the Commission did not argue the Allied Signal-Honeywell merger had led to widespread bundling. From the Commission’s perspective this was only because the merger was too recent and thus the impact of bundling had not yet been felt.

EVIDENCE THAT BUNDLING WOULD LEAD TO EXIT BY RIVALS

At the same time that the Commission was arguing that the Honeywell-Allied Signal bundling effects were slow to arise, the Commission took the stance that the GE-Honeywell merger would lead to nearly immediate exit or marginalization of rivals.

Is exit (or marginalization) really likely in this business? Here the Commission seemed to rely on the dire warnings of some competitors. Of course, these competitors are not disinterested parties. They would hope to

“If the exit or marginalization were to occur more gradually, then one would want to take into account the benefits gained by customers during the period of lower prices. Such a calculation was never made.
block the merger or purchase various Honeywell "jewels" that the commission would require to be spun off.

Nor was there evidence provided that the rival firms were in any danger of exit. To the contrary, evidence was presented that showed the long-term viability of the rival firms. For example, the stock market’s reaction did not anticipate the financial vulnerability of aerospace firms. From the time of the announced merger to the time of the EC hearing, almost all of the rival firms had a gain in stock price that exceeded the S&P 500 index. The firm that underperformed the index was GE.

One would not expect to see aerospace firms concede defeat quickly. Airframes are long-lived. The typical plane is on the market for twenty-five or more years. Thus, a contract gained today (or ten years ago) would provide a long stream of profits down the road. Hence, even if firms were precluded from new contracts, they are not going to disappear any time soon. In fact, the next plane for which the engine choice would be made was the Boeing 787.

In the aerospace industry there are several large players who have an interest in maintaining competition. Military purchases play this role. Both Rolls Royce and P&W have multibillion-dollar annual revenues from military contracts, and there is a substantial amount of spillover between civilian and military work. While it could be argued that no single airline would be eager to provide the public good of maintaining competition, Airbus and Boeing (each of which has about a 50 percent market share and each of which works closely with airlines when making purchase decisions for the aircraft it designs) do have this incentive.

**POLICY PRESCRIPTIONS**

We have discussed at some length whether competition authorities should be worried about mixed bundling in the context of the GE-Honeywell merger. Putting relevancy to GE-Honeywell aside for a moment, in an environment in which the Cournot effect does exist, there remains the issue of what should be the antitrust policy toward a merger of two producers of complements, each with market power. Another way of saying this is that if we did indeed find a case where the Cournot effect was large, what should we do about it?

At first blush, the answer would seem to be nothing. Prices fall! Social welfare is higher. Consumers gain, and competitors lose. This is not the arena for antitrust authorities to get involved.”

“While the European law is concerned with market dominance, the economic rationale would be a concern about high prices. Here the problem is with low prices. The European law does not seem to take efficiencies into account that might lead to lower prices (Schmitz 2002).
The European Commission took a novel perspective. It was concerned that the long-run impact of the combination could be to put competitors out of business. Similar to a predatory pricing case, once the competitors were either hobbled or vanquished, the merging firm would have even more power to raise prices, and then social welfare would fall.

Unlike in the typical predatory pricing case, recoupment might or might not be an issue. Depending on the parameters and setup of the model, the combining firms might actually make money in the process of putting their rivals out of the market.

If we follow this line of reasoning, a several-part-test will follow. This test is an extension of one developed by Carl Shapiro and by the Department of Justice."

1. Is there an incentive to bundle?
   a) Under what circumstances does the combined firm earn higher profits through a bundled pricing strategy?
   b) Did either firm have an opportunity to bundle prior to the combination? If so, is there evidence that bundling is a common practice in this industry?
      i) If we see bundling, then what is the marginal impact of increasing the potential scope of the bundle?
      ii) If we do not see bundling, then how do the opportunities created by this combination create a different incentive to bundle?

2. What is the immediate gain to consumers from lower prices?
   a) How much do we expect prices to fall due to bundling?

3. What will be the impact on competitors?
   a) How much will competitors’ price fall?
   b) What will be the shift in share?

4. How long do we expect these lower prices to persist?
   a) How long do we expect the rivals will be able to hold out?
   b) Are rivals sufficiently close to exiting the market that this will tip the scales?
   c) Are there large customers with market power that have an incentive to keep multiple firms in the market?

"The Shapiro test was presented to the Merger Task Force. The DOJ guidelines are described in a report submitted to the OECD roundtable on portfolio effects in conglomerate mergers, DOJ (2001)."
5. If the rivals exit, what is the expected harm?
   a) Will other firms be able to enter the market?
   b) Will large buyers be able to hold prices down?
   c) Or, if prices rise, what is the expected damage?

In short, there are immediate benefits to the combination. How long can we expect these benefits to last? How likely is the potential harm, and how big is the potential harm? Or, even simpler, what is the present value of the net expected change in social welfare from allowing the combination?

Applying this approach to the GE-Honeywell case, the Commission stopped at step 1. Even if bundling were to threaten the long-term viability of competitors, no analysis was made of the trade-off. Properly discounted for time and the fact that the outcome is uncertain, was there any argument to be made that the net present discounted value would be negative? The Commission made a recommendation against a merger that by its own account would be expected to lower prices without demonstrating that the expected long-run harm would outweigh short-term gains.

**REMEDIES**

When the net impact of some conduct is negative, the next step is to look for remedies, either structural or behavioral, that would solve the problem. In the case of bundling a very simple behavioral remedy presents itself. The firm can commit itself not to bundle.

Since it was the potential to engage in this behavior that concerned the Commission, the fact that the merging parties would agree not to offer bundle discounts should signal that this was not an important element of their incentive to merge. And since it was the bundling behavior that had the potential to lead to dominance, without bundling there would be no dominance.

A no-bundle discount policy is straightforward. A firm can do that by having to provide an itemized breakdown of the package price and giving a price to each component in the package. The individual prices must add up to no more than the bundle price.

A bundle works only if it is offered at a discount to the components. A firm need not be regulated on what it can charge for each item. It can charge whatever it would like for components and for the bundle just so long as the combined price of the components is not more than the bundle.

The Commission rejected this approach based on its preference for structural over behavioral remedies. Yet, this particular behavioral solution does not appear hard to monitor or enforce. If the component elements of a firm’s bids are not itemized or add up to more than the package total, that would be an automatic violation.
CONCLUSION

In the Statement of Objections, the Commission presented a theory of bundling that was based on the premise that the merged firm would be able to bundle and would have a rational economic incentive to do so. The result of this bundling would be prices so low that competitors will be foreclosed.

When closer scrutiny was given to this argument, it turned out not to apply to the market for aircraft engines, avionics, and nonavionics products. The Commission subsequently abandoned its original approach:

The various economic analyses have been subject to theoretical controversy, in particular, as far as the economic model of mixed bundling, prepared by one of the third parties, is concerned. However, the Commission does not consider the reliance on one or the other model necessary for the conclusion that packaged deals that the merged entity will be in a position to offer will foreclose competition from the engines and avionics/non-avionics markets.

Instead, it based its decision on a new, dynamic, theory: Foreclosure of competitors would occur as a result of predation accomplished through the cross-subsidization of bundled sales. As explained in the EU’s Competition Policy Newsletter (Giotakos 2001):

Thanks to GE’s strong generation of cash flows resulting from the conglomerate’s leading positions on several markets, following the merger, Honeywell would have been in a position to benefit from GE’s financing surface and ability to cross-subsidise its different business segments, including the ability to engage in predatory behaviour.

There may be a simple explanation as to why reliance on the Rolls Royce model was so attractive to the Commission. If, in fact, a merged firm would make more money by reducing prices when selling complementary products, then there is no need to estimate the cost of predation (as there is none), and there is no need to estimate recoupment (as none is required). Moreover, the merger becomes the proximate cause of the price discounting, and so there is now a reason to think that this type of economically rational predation will occur postmerger even if it does not occur premerger.

Against this background, the Commission may have been disappointed to discover that the economic models upon which it had relied—the model advanced by Rolls Royce, and even the basic Cournot complements model—were unsuited to the task.

32 Paragraph 352 of the Decision.
33 Predation was never seriously discussed in the Statement of Objections: There is only a single mention of predation in the SO, unrelated to the bundling theory, and this occurs in a footnote.
34 The Cournot complements model has a crucial assumption, namely that each firm charges one price in the market to all of its customers. While this may be an acceptable assumption for a consumer good, it does not apply to the negotiated sales of aircraft engines and avionics. With perfect
Although the Commission abandoned the original model, it did not replace the flawed model with another model. Rather, the Commission switched to a dynamic theory of predation. But it never carried out any of the steps required to establish the facts necessary to support that theory.

In the end, the European Union Merger Task Force did back away from its economic theory of bundling. But it did not back down from its conclusion that bundling was a reason to block the merger. The decision in this case presents a challenge for the role of economic analysis in the design and implementation of antitrust policy.

POSTSCRIPT

General Electric and Honeywell each appealed the decision to the European Court of First Instance (CFI). On December 14, 2005, the appeal was denied.

The CFI rejected the primary arguments given by the European Commission to prohibit the merger. The court was not persuaded that GE and Honeywell would engage in bundling. Nor was it persuaded that GE would leverage the financial strength of its leasing business (GECAS). The CFI also rebuked the Commission's conclusion that GE would start to engage in anticompetitive conduct based solely on its greater vertical and horizontal scope in the market.

While the primary economic arguments against the merger were rejected, the CFI upheld the Commission's decision on what could best be described as a series of technicalities. The court pointed out that the merger would create market concentration in three markets: (1) small marine gas turbines, (2) engines for large regional jets, and (3) corporate jet aircraft engines. Since these issues had not been resolved, the merger could be blocked.

Even if it were true that the small marine engine business might have been concentrated, this denies the nature of the merger discussions with the European Commission. One can safely bet that the leadership of GE and Honeywell would have been willing to divest the small marine engine business as a condition of the merger. Similar solutions could have been found for large regional jet engines and corporate jet engines. The problem was that the negotiations with then-Commissioner Mario Monti broke down over the major issues of GECAS and bundling, leaving the tertiary issues

information, there would be no Cournot complements effect at all. Even with imperfect information, the Cournot effect is diminished to the extent that firms can differentiate between their customers. The evidence suggests that prices and profit margins vary substantially among customers. But the Commission did not attempt to estimate the degree of uncertainty and distribution of preferences so as to employ the negotiation model presented to it.

“The category of large regional jet engines was the most significant of the horizontal areas. The overlap involved indirect competition at the aircraft level (not direct competition at the engine level) relating to a single niche aircraft, the Avro. The Avro was cancelled in the fall of 2001 after 9/11 sent the aviation business into a nosedive.

Page 439
such as overlap in marine and regional jet engines unresolved. Of course, even if the CFI had overturned the commission, it was too late to resurrect the merger.

REFERENCES


INTRODUCTION

During the early 1990s Toys 'R' Us (hereafter, TRU), by a considerable stretch the largest retailer of toys in the United States, persuaded leading toy manufacturers to restrict the range of products they sold to warehouse clubs—a newly emerging form of consumer goods retailing. The restrictions were challenged under the antitrust laws by the Federal Trade Commission (FTC) in May 1996. The extensive litigation that followed clarified in important ways the kinds of "vertical restraints" a retailer could impose upon its suppliers and the role of efficiency defenses, and especially claims of free-riding on services provided by the retailer, in justifying such restraints.

INNOVATION IN RETAILING

The tensions that arose among TRU, warehouse clubs, and toy manufacturers during the 1990s were a natural sequitur from the several revolutions that occurred in retailing during the prior century and one-half (Chandler 1977, chs. 1 and 7; Adelman 1959; and Scherer 1999).

As the United States emerged from its great Civil War, most consumer goods were provided at retail by general stores or "mom and pop" stores.

*The author was testifying expert economist on behalf of the Federal Trade Commission in this matter.
'The company’s logo prints the “R” backward, which is not easily accomplished in type-set publications. We use quotation marks to note that our usage is not strictly accurate.

In the matter of Toys “R” Us, Inc., docket no. 9278.
These outlets were challenged competitively at first in the larger cities by department stores and then by mail-order houses such as Sears, Roebuck. During the twentieth century new retailing approaches appeared successively in the form of the chain stores (such as A&P and Walgreen’s), supermarkets, hypermarket chains (such as Kmart and Wal-Mart), and specialized “category killer” chains (such as Home Depot for hardware and garden supplies and Staples for office supplies). These innovations in retailing were often accompanied by reductions in the percentage margin between prices and wholesale merchandise acquisition costs (hereafter, percentage retail margin, or PRM), and hence reductions in the prices that consumers paid, to the competitive disadvantage of more traditional retailing forms.

The traditional retailers fought back in part by seeking governmental protection—for example, in unsuccessful attempts during the 1880s and 1890s to secure state legislation curbing department stores’ lower prices, through delays in the spread of free rural mail delivery aimed at inhibiting the growth of Sears, Roebuck, and in the passage of tougher antiprice discrimination laws (the Robinson-Patman Act) and laws enabling mandatory minimum resale price maintenance (so-called fair trade laws) during the 1930s. Gradually, these barriers proved ineffective or were repealed, and the retailing innovations took root and spread, with consumers as principal beneficiaries.

Toys “R” Us made important contributions to this innovative process. Founded by Charles Lazarus in 1948, with expansion to a Washington, D.C., branch first bearing the Toys “R” Us logo in 1954, it was one of the first category killer discount chains. It offered an unprecedentedly broad line of toys, encompassing some sixteen thousand different items (stock-keeping units, or SKUs) by the 1990s (reduced to roughly eleven thousand in 1996), and priced its products to realize margins between selling price and acquisition cost well below the 40 to 50 percent values that were the norm in more traditional toy outlets. It reached the fifty-store threshold in 1974 and continued its expansion to operate 497 retail toy outlets in the United States during 1992, along with 126 stores in other nations, to which it had introduced the previously unfamiliar concept of systematically discounted toy prices.

The Rise of Warehouse Clubs

The warehouse clubs, among which the most important examples were Costco, the Price Club, Sam’s (affiliated with Wal-Mart), Pace (affiliated with Kmart), and BJ’s, made further additions to retailing innovation history. At first, beginning in 1976, they sold merchandise mainly to small business customers. But in the late 1980s they began accepting individual retail customers who paid an annual membership fee of roughly $30, or approximately 2 percent of the average member’s annual purchases. They purchased selected nationally branded food, appliance, electronic,
automotive, and other consumer products in large quantities, transported them to austere warehouse-like buildings at low-rent locations, and moved them onto the selling floor still heaped upon the pallets with which they had been transported from the manufacturer to the store. Customers served themselves from the pallets and packed their purchases into oversize shopping carts after bringing them through a streamlined checkout counter. Despite selling a wide array of consumer products, the typical warehouse club outlet carried only some three thousand individual SKUs, with the product mix varying over the year to reflect both seasonal demands and the opportunistic purchases made by the clubs' buyers. Because the product mix varied so widely, shopping at the clubs was viewed by many consumer members as a kind of treasure hunt—one never knew what specials one might encounter.

Because of their extraordinarily low overhead, the large volumes in which members typically bought, and the annual admission fee, the clubs were able to offer their goods at uniformly low percentage retail margins—on the order of 9 to 12 percent. By 1992, there were 576 warehouse club stores in the United States. Their appeal was spreading rapidly, and marketing researchers projected in 1992 that the number of club outlets would double during the following decade.

The low PRMs taken by the clubs on the 100 to 250 toy items they stocked, depending upon availability and season, were perceived as a serious threat by the management of TRU. Although its principal draw to consumers was the huge array of products TRU stocked, it also emphasized its role as a low-price, if not the low-price, outlet, leading consumers to sample that array. To be sure, it faced price competition from other toy retailers—increasingly, from hypermarkets led by Wal-Mart, which was expanding rapidly from its Arkansas base during the 1980s and 1990s, and which carried as many as three thousand toy items at seasonal peaks, sold at PRMs averaging 22 percent but considerably lower for "hit" items.

TRU responded to this competition by varying its prices inversely with the popularity of the item, and hence with the extent to which consumers might make price comparisons at alternative outlets. The best-selling hundred or so TRU toys were sold at substantially discounted PRMs; margins then ascended into the 32-38 percent range for products ranked 1000 to 10,000 in sales volume. Especially after competition from Wal-Mart intensified, TRU recognized that it could not always be the lowest-price outlet, but it tried to keep prices of the most popular items close to those of the toughest rivals. Customers were drawn to TRU through the advertising of a few hundred low-margin, popular items (along with disposable diapers and baby formula, priced at razor-thin margins, and shelved near the back of the stores); TRU's profits were then enhanced substantially as the typical customer made several impulse purchases of higher-margin products.

With their uniform low-margin policies, the clubs added a new quantitative dimension to this marketing equation. The prices they asked on popular items were seen by TRU executives as a serious threat to TRU’s reputation as a low-price vendor, and hence to TRU’s ability to draw customers in for additional higher-margin purchases. To be sure, this adverse “image” effect could be minimized by reducing TRU’s prices on the relatively few popular products stocked by the clubs. TRU’s initial reaction to the clubs’ rapid ascent in the early 1990s was to make selective downward price adjustments, with a maximum negative impact on its retail margins estimated at roughly $55 million per year. But it sought a better solution.

Following preliminary discussions with many of the leading toy manufacturers, TRU made the club problem the main theme of its meetings with toy manufacturers at the New York Toy Fair in February 1992—the annual meeting at which retailers join manufacturers to scrutinize the newest toy designs and book orders for the forthcoming pre-Christmas peak season. Through these meetings, TRU articulated to toy makers what its policy would be toward manufacturer sales to warehouse clubs:

- Toy makers should sell to the clubs no new or advertised products unless the clubs purchased the manufacturer’s entire product line.
- All special products, exclusives, and close-out clearance items offered to the clubs should be shown to TRU first to see whether TRU wished to preempt the clubs and stock them itself.
- Old and basic products should be sold in special combination packages, for example, with a bridesmaid doll added to a basic Barbie Bride doll package.
- There was to be no discussion between TRU and the toy makers as to the prices at which these special, typically more complex and costly, products were sold to the clubs.

The penalty for violation was that TRU would not purchase offending products.

Subsequent fine-tuning discussions clarified the basic theme: that the manufacturers would not sell to the clubs products that were identical to those stocked by TRU, rendering it difficult for consumers to make straightforward comparisons between club and TRU prices. Since TRU tended to order all products expected to become best-sellers or receive television advertising support, this meant that the clubs would be able to obtain only by happenstance the most popular “hit” merchandise. (Predicting what new toys will catch consumers’ fancy is an extremely uncertain game; even the most experienced retailer executives err.)

Most of the leading manufacturers accepted TRU's new policy with minor variations, and the supply of best-sellers to clubs atrophied. Despite the continuing growth of club merchandise sales more generally and an increase in the number of club outlets to 695 in 1995, warehouse clubs' toy sales peaked at 1.9 percent of total U.S. toy sales in 1992 and then declined to 1.4 percent in 1995. By 1993, TRU no longer found it necessary to adjust "high profile" products' prices downward to deal with club competition.

THE ANTITRUST CHALLENGE

The clubs responded to these changes in part by threatening legal action against TRU and the manufacturers, but they also informed the FTC of what had happened. Following an investigation, the FTC, pursuing more activist policies after twelve years of relative inactivity during the Reagan and Bush administrations, issued in May 1996 a formal complaint charging TRU with diverse antitrust violations.

Critics of U.S. antitrust policy often assert that in responding to information provided by aggrieved rivals (e.g., the clubs, as rivals of TRU), the federal antitrust agencies are protecting "competitors" rather than the processes of competition and hence consumer interests, as should properly be their mandate. There are two responses to this criticism. First, antitrust agency staff are acutely aware of the criticism and try hard to limit their cases to those that serve the broad public interest. Whether they succeeded in the TRU case is for the reader to judge. Second, agency staff have only limited insight into what is happening in the vast world of business enterprise. Consumers, too, and especially consumers at retail, lack the information needed to defend themselves against abuses. Without information from aggrieved market participants and other affected parties, the antitrust enforcement agencies would overlook many genuine violations of the antitrust laws.

The complaint against TRU was the subject of a bitterly contested trial before an FTC administrative law judge, who found in September 1997 that the antitrust laws had in fact been violated. As is customary, the judge’s decision was appealed by the losing party, in this instance TRU, to the FTC as a whole. In a meticulously reasoned decision that was plainly intended to make a definitive pronouncement on the extent to which retailers could initiate restraints limiting the purchasing opportunities of their rivals, the four Commission members (one of the five commissioner slots was vacant at the time) affirmed that TRU had violated the antitrust laws.

'Opinion of the Commission, 126 FTC 415, 597 and note 15 of the Commission's opinion. 
'In the Matter of Toys "R" Us, initial decision, September 25, 1997, 126 FTC 415, 418.
The Vertical Restraints

TRU was charged with, and found responsible for, two rather different violations of the antitrust laws. For one, the agreements it elicited from individual toy manufacturers not to sell popular items to the warehouse clubs fell under the rubric of vertical restraints— that is, where a firm (i.e., a retailer) located at one tier in the chain from production to consumption imposes restraints upon the distribution practices of firms (in this case, manufacturers) located at a different tier in the chain. But in addition, TRU was accused of orchestrating a set of horizontal agreements among manufacturers to deny popular products to—that is, to boycott—the warehouse clubs.

Antitrust law makes a distinction between unilateral vertical restraints— for example, TRU announces, "If you sell product X to the clubs, we will not stock it, end of discussion, full stop"—and restraints between vertically stacked parties in the distribution chain that comprise a meeting of minds and hence mutual agreement between the parties. The relevant case precedents show greater willingness to accept vertical restraints imposed unilaterally, as compared with those achieved by agreement, without a detailed investigation of the pros and cons.

Although TRU claimed that in announcing to toy manufacturers its intent not to stock products they sold to the warehouse clubs it imposed only unilateral restraints, the facts clearly established that what occurred was much closer to a bilateral meeting of minds. TRU executives met repeatedly with each major toy manufacturer to present their list of demands. There was give-and-take in the discussions to modify the policies and adapt them to special circumstances. TRU sought, and from at least ten leading manufacturers received, assurances that the manufacturers would conform to the negotiated restrictions on supply to clubs. TRU executives inspected in advance the special products toy makers proposed to sell to the clubs and in some instances requested, and obtained, modifications that differentiated them more clearly from products stocked by TRU. There was continuing feedback as TRU monitored the availability of manufacturers' products on the shelves of warehouse clubs and reported apparent policy violations to the manufacturers, threatened and in some cases removed offending products from TRU shelves (actions that standing alone would be unilateral), and through subsequent negotiations achieved policy convergence.

Considering this record, the Commission found that a series of anticompetitive vertical agreements was reached that prima facie violated Sherman Act Section 1, but whose legality depended upon the analysis of additional considerations, to be addressed here shortly.

The precedents are discussed at length in the Opinion of the Commission, 126 FTC 415, 569-615.
The Horizontal Boycott

Three of the four FTC commissioners, with one dissenting, concluded that TRU also acted as the "hub" in a horizontal "hub and spoke" agreement among toy manufacturers to deny top-line and advertised products to the warehouse clubs. Such horizontal agreements are treated harshly under the antitrust laws as illegal per se, with consideration of extenuating circumstances only when the agreements yield plausible efficiency gains that cannot be achieved through alternative policies restricting competition less—conditions that are difficult to satisfy.

What gave rise to a horizontal agreement problem was the fact that most leading toy manufacturers viewed the clubs as an attractive and rapidly growing outlet for their products, sales to which would among other things lessen their growing dependence upon TRU as their largest and most powerful customer. To placate TRU and prevent it from taking product placement actions detrimental to their interests, the toy makers were willing to go along with TRU's proposed restraints—but only if they could be assured that they were not, in so doing, sacrificing sales to rival manufacturers.

Toys are highly differentiated products; as a rule, one toy maker does not consider itself to be competing head-to-head with all other toy makers, but rather with only a handful of firms that offer products (especially advertised products) similar in function and design to its own products. The manufacturers expressed concerns to TRU that if they complied with the proposed TRU policies, specifically named rivals might not, causing them to sacrifice significant sales in the clubs. As a top TRU executive testified, "They would always tell us, 'I'm only there because my competitor is there.' And we would say, 'Well, he keeps saying he's only there because you're there.'" To deal with the problem, he testified, "We communicated to our vendors that we were communicating with all our key suppliers. . . . We made a point to tell each of the vendors that we spoke to that we would be talking to other key suppliers."

TRU executives repeatedly informed toy manufacturers that a "level playing field" was being maintained and that their rivals had agreed to pursue the TRU-suggested policies. They received manufacturer complaints about rival noncompliance and communicated back to the originators reassurances that perceived deviations had been eliminated. And this, three of the four FTC commissioners concluded, was a classic "hub and spoke" horizontal agreement. The dissenting commissioner questioned the strength of the factual evidence supporting a horizontal agreement inference and argued that "TRU's very indispensability gave each toy manufacturer every

"Testimony of Roger Goddu, quoted in Opinion of the Commission, 126 FTC 415, 554.
"Ibid, at p. 55, from CX 1658 p. 278.
"Among the precedents cited was Interstate Circuit Inc. v. U.S., 306 U.S. 208 (1939)."
incentive—every unilateral incentive—to knuckle under to TRU’s demands regarding the clubs,” rendering horizontal agreements among the manufacturers unnecessary."

A limitation of both the horizontal and vertical agreement allegations is that the agreements were not pervasive among toy manufacturers. The vertical agreements were found to have been implemented only by ten named producers, the horizontal agreements by seven—to be sure, those who originated most of the nationally advertised toy products." Electronic game maker Nintendo in particular rejected TRU’s overtures, in part because it sold its products through electronic specialty stores and was therefore less dependent upon TRU. Nintendo’s noncompliance in turn posed problems for rival Sega, which adhered only intermittently to TRU’s proposed policy. And at times Little Tikes, a manufacturer of large blow-molded toys, opted out because compliance might jeopardize more voluminous sales to warehouse clubs by its parent, the Rubbermaid Corporation. But there are strong antitrust law precedents holding that substantial but less than complete compliance with restrictive agreements does not reverse the illegality of such agreements if the impact on consumers is substantial.

The Role of Market Power

An appreciable fraction of the testimony by economists in the TRU case addressed the question of whether TRU possessed “market power” in the sense defined by prior antitrust precedents. As events ensued, this emphasis was misdirected, since proof of market power is normally unnecessary to infer illegality when horizontal agreements are shown. But for the vertical aspects of the case, the existence of market power would strengthen an inference of illegality, since an actor at one stage in the vertical chain of distribution was more likely to implement restraints with a meaningful impact on consumers if the initiator at another stage had a powerful position, and significant anticompetitive effects would ensue if the restraints applied to substantial shares of the affected markets.

The concept of “market power,” said by a reviewing appellate court to imply a degree of market dominance less than “monopoly” in the sense normally used by economists, is not well defined. There are at least three ways of showing that it exists or does not exist—structure-oriented measurement of market shares in relevant markets, statistical analyses of the relationship between prices and market shares, and analysis of the effects of restrictive actions on market outcomes. All three avenues were pursued in the Toys "R" Us case.

"Opinion of Commissioner Orson Swindle dissenting in part and concurring in part, 126 FTC 415, 620 (emphasis in original).

″Opinion of the Commission, 126 FTC 415, 575.

Although there are hundreds of toy manufacturers and thousands of firms retailing toys in the United States, both markets exhibit what at face value would be called intermediate concentration levels. At the manufacturing stage, the top four U.S. suppliers during the early 1990s originated from 34 to 45 percent of total toy supplies (many imported from southeast Asia), varying with whether only traditional or also video game toys are counted and to some extent with the data source used. However, the focus of TRU’s clubs policy was on nationally advertised toys, and for those, roughly two-thirds of the national television advertising was done by only eight toy makers, three of them divisions of Mattel, the source inter alia of Barbie dolls. By accepted antitrust standards, toy manufacturing was sufficiently highly concentrated that an inference of “market power” could be supported.

At the retailing level, Toys “R” Us accounted for approximately 20 percent of all retail toy sales in the United States. However, all parties acknowledged that the relevant retail markets for toys were localized. TRU maintained stores principally in the larger metropolitan areas. Its marketing research revealed that in population areas within a thirty-minute drive from a TRU store, its average share of the total toy market was approximately 32 percent—again, a structural position sufficient to imply the existence of “market power” under the received vertical restraints precedents.

TRU maintained rich internal data on the competitive structure of the localized retail markets surrounding its U.S. sales outlets. Tapping these data and information on individual store percentage retail margins, economists representing TRU computed regression equations relating store PRMs to the number of significant rivals (mainly, hypermarkets such as Wal-Mart and Target as well as, briefly during the early 1990s, warehouse clubs) within a local market. They showed that TRU realized somewhat lower PRMs in markets with one or two major rivals than in markets where no major retailing rival existed, and that the presence of additional rivals beyond two made no significant difference. The PRM differences with and without major competition were said to be sufficiently small as to be de minimus. However, TRU adjusted prices to deal with local market conditions for only about 250 items, typically those that were nationally advertised and/or “hot,” out of the nearly sixteen thousand items it stocked at the time. For the majority of SKUs, prices were uniform nationally. A witness for the FTC testified that, given the modest share of total TRU sales associated with the items whose prices were geared to local competitive conditions, the implied price differences on those competition-sensitive items could be as much as 6 to 7 percent—a not negligible magnitude.

Economists representing TRU argued, using regression and qualitative evidence, that TRU’s ability to raise prices in any given retail market was severely constrained by the presence of vigorous low-price competitors such as Wal-Mart, Kmart, and Target, and therefore that TRU lacked the “market power” required to find its vertical restraints inconsistent with antitrust law. It is undoubtedly true that TRU’s ability to raise prices was
constrained by strong competitors. However, the argument skirted a fundamental point. The purpose of TRU’s warehouse clubs policy was directed not toward raising prices in local retail markets whose structure reflected the encroachment of hypermarkets, but toward curbing additional and rising competition from warehouse clubs that could otherwise force reductions in retail prices. To the extent that TRU’s warehouse club policies were successful in suppressing the clubs’ competitive threat, downward price changes were avoided—a clear manifestation of market power. As the FTC concluded, rejecting the constrained price increase argument, “there is little question that the boycott of the warehouse clubs that TRU organized could and did lower output by avoiding a decrease in toy prices by TRU and TRU’s non-club competitors.”

**The "Free-Riding" Efficiency Defense**

The U.S. courts have accepted as a defense to some vertical restraints the argument that such restraints help solve problems that would otherwise render the distribution of goods or services less efficient. The basic theoretical premise was formulated by Telser (1960). Telser observed that retailers often provide présale services that help consumers make well-informed product choices and, by enhancing the demand for the products, are valuable to manufacturers. Although other présale services might fall within the scope of Telser’s argument, the standard example is a retailer that maintains display models in its showroom, explains to consumers (or allows them to experiment and learn) how the product functions, and hence enlightens their choices. Providing these services usually entails costs for the retailer. However, consumers might visit the high-service retailer’s showroom, utilize the présale services, and then travel down the street to a low-service outlet and buy the product at discounted prices. In this case, the discount retailer is said to "free-ride" on services provided by the high-service retailer, and if such free-riding occurs with sufficient frequency, the high-service retailer will lose its incentive to provide desirable services and the market will implode to a low-service equilibrium. Restricting the availability of products to low-service retailers and/or preventing them from quoting steeply discounted prices restores the incentive for retailers to provide the desired présale services.

How important such free-riding market failures are in the real world, their implications for consumer welfare, and whether there exist alternative, less restrictive ways of solving the free-riding problem are questions on which unusually intense disagreement exists among professional economists, with those advocating vertical restraints identified more or less closely with the so-called Chicago School of thought."

"Opinion of the Commission, 126 FTC 415, 597 (emphasis in original).

For sharply differing textbook views on the problem, compare Carlton and Perloff (2000, ch. 12); and Scherer and Ross (1990, ch. 15).
There was little dispute over the fact that TRU provided valuable présale services. With its large selection of products, it acted as a kind of “showroom” for toy industry manufacturers, ensuring that most of their products, and not just the best-sellers, were available to consumers. It tended to order products for the Christmas rush a few weeks earlier than did other toy retailers and to take delivery of the products slightly earlier, helping the manufacturers economize by spreading their production over a longer time period. It claimed also that its early stocking decisions signaled to other retailers what products were likely to be in strong demand, but this was contradicted by evidence showing that other retailers, and especially the warehouse clubs, made purchasing decisions on the basis of their own independent market assessments, taking into account manufacturers’ announced advertising plans. Perhaps most importantly, several times a year TRU placed in leading metropolitan area newspapers catalogue inserts with full-color illustrations of several hundred products featured on its shelves, informing consumers of what was available and at what prices.

There were, however, several logical and factual problems with the TRU free-riding argument. For one, TRU’s large product selection was unlikely to have provided présale information on which consumers then free-rode to make purchases at warehouse clubs. TRU merely put its products on its shelves, without providing the kind of demonstrative services emphasized in the original free-rider theories. The average price of items advertised in TRU’s newspaper insert catalogues in the spring of 1997 was $45.41, the median price $29.99. With such low ticket prices, few consumers inspect the product in a TRU showroom and then make a special trip to buy the item at a lower-price (e.g., warehouse) outlet. The FTC found no evidence that consumers “sought demonstration or explanation of a toy at TRU and then purchased the product at a club.” In addition, TRU derived direct competitive advantage in consumers’ eyes from having the largest selection of toys, and thus its large stock conferred consumer image benefits upon itself and was not something on which rivals plausibly free-rode.

Second, the costs that TRU incurred to take early delivery of seasonal products and to inform consumers about available toys and their prices through its widely distributed catalogues and newspaper inserts turned out not to be very high. It was compensated for its early stocking practices by the deferral of invoice payment requirements—for example, from June, when delivery was taken, to December, when sales were at peak levels. When products taken into inventory turned out to be “duds” in consumers’ eyes, manufacturers compensated TRU for the ensuing closeout discounts by granting substantial retroactive wholesale price discounts. Indeed, manufacturers testified that no other retailer received deferred payment terms and closeout discounts as favorable as those accorded TRU. They also paid

“Opinion of the Commission, 126 FTC 415, 603-604.
TRU advertising allowances to compensate for the cost of distributing catalogues to consumers. The case evidence revealed that more than 90 percent of TRU's advertising outlays during the mid-1990s were reimbursed by manufacturers. If TRU bore little of the cost of providing presale advertising and early product stocking, it was implausible to infer that free-riding by others on those services could lead it to eliminate them, as the standard theory of free-riding assumes.

TRU's principal economic witness attempted to demonstrate through a series of regression analyses that other retailers nevertheless experienced sales increases as a result of TRU's catalogue advertising, and hence free-ride on that advertising. What ensued was a war of alternative regression equation specifications. The first TRU analysis revealed that for items featured in TRU's April 2, 1995, catalogue, TRU achieved substantial sales gains relative to nonadvertised items. But in addition, other toy retailers experienced smaller but still appreciable sales gains on those items relative to nonadvertised items. The FTC's economic expert questioned whether there might have been something special about the items selected for the April 1995 catalogue that induced high subsequent sales growth independent of the advertising effect—an example of what econometricians call "omitted variables" bias. Re-estimating the original TRU regression equation with variables added to measure the pre-April growth momentum achieved by advertised as compared to nonadvertised products, the FTC analysis showed that catalogue inclusion indeed augmented TRU's sales, but had a negative impact on rival retailers' sales of the advertised products.

Another round of competing regression equation specifications yielded similar inference reversals, prompting the administrative law judge to call a halt to the war and view the statistical inferences as not proven.

Finally, in rejecting TRU's free-rider defense, the FTC observed that before February 1992, when the warehouse club policy was announced at a Toy Fair, "no toy company document... even hints that 'free-riding' by one toy retailer on the efforts of another could be a problem in the industry." Rather, TRU's concern was the damage that low warehouse club prices could do to its reputation as a low-price merchandiser and the necessity of reducing popular toys' prices to warehouse club levels to avert that adverse reputation effect. The first recorded mention of free-riding in the voluminous case record occurred later in 1992, when the clubs threatened to sue TRU and its suppliers for discriminatory sales policies. The record is silent as to whether the notion originated spontaneously or whether it was suggested by outside consultants hired to evaluate the legal threats. Whatever the origin, the Commission concluded that TRU's concerns about

---

"National television advertising was sponsored almost exclusively by the manufacturers, not TRU. Toy advertising outlays on television in 1995 amounted to roughly five times TRU's outlays for mostly local catalogue advertising.

"Opinion of the Commission, 126 FTC 415, 567. See also pp. 579-580 of the opinion."
free-riding as a justification for its restrictive policies were a "pretext" and not a valid basis for the policies' implementation."

THE OUTCOME

Rejecting the free-rider defense and concluding that TRU had orchestrated both a horizontal boycott and restrictive vertical agreements, the FTC ruled that TRU had violated the antitrust laws. It ordered TRU to cease entering into vertical agreements with its toy suppliers that limited the supply of toys to discounters, to cease facilitating horizontal agreements among its suppliers concerning the sale of toys to other retailers, and to refrain from requesting information from manufacturers about their sales to toy discounters.

Throughout the proceedings, counsel for TRU openly expressed doubts that they could obtain a fair trial at the FTC, combining as it does in a carefully compartmentalized way the functions of prosecutor and judge. TRU therefore sought, as many losing parties do, reversal from an independent federal appellate court. Operating in all parts of the United States, it could as the appealing party choose the jurisdiction in which to file its appeal. It selected the Seventh Circuit Court of Appeals, situated in Chicago, with three members of the University of Chicago law faculty as sitting judges. It presumably believed that its chances were best in the environment where the theory of free-riding had originated. If so, then it was mistaken.

A panel of three appellate judges headed by a University of Chicago law faculty member unanimously ratified the FTC's decision. It found that the horizontal boycott, to which the contested market power issues were less relevant, clearly violated Section 1 of the Sherman Act. Observing that, given this conclusion, it did not have to rule formally on the case's vertical issues, it nevertheless opined that TRU's pleadings "fundamentally misunderstood the theory of free riding." This was so, the court said, because in the absence of TRU threats, most leading manufacturers wanted to sell to the clubs without restrictions and since, with manufacturers paying for most of TRU's présale services, there was little opportunity for meaningful free-riding.

THE AFTERMATH

The Federal Trade Commission's October 1998 decision finding TRU in violation of the antitrust laws precipitated a flurry of additional class action

"Opinion of the Commission, 126 FTC 415, 607.
antitrust suits against TRU and leading toy manufacturers, some by private parties seeking compensation for damages sustained and one by the attorneys general of forty-four U.S. states plus Puerto Rico and the District of Columbia. “To avoid the cost and uncertainty of protracted litigation,” TRU along with three toy manufacturers negotiated in 1998 and 1999 settlements valued at a total of $56 million, some paid in cash and most in the form of toys to be distributed to needy children.

In addition to the rejection of its policies toward warehouse clubs, whose overall sales growth slowed during the late 1990s, TRU experienced increasing competition from hypermarkets, and especially from a Wal-Mart that was continuing its “march to the sea.” Wal-Mart’s toy sales share exceeded TRU’s beginning in 1999 and continued rising to nearly 30 percent of national toy sales by 2006. TRU’s share declined to an estimated 15 percent in 2006, when TRU was overtaken also by Target. TRU’s profits ebbed and turned negative in some years, despite strenuous efforts to modernize many stores, close others, secure exclusive deals on toys that were predicted to become hits, and induce its staff members to devote more of their time helping shoppers.

With its stock price stalled at only 60 percent of its all-time peak value, TRU agree in 2005 to “go private” in a $6.6 billion transaction. Two other toy retailing chains, KB Toys and F.A.O. Schwartz, filed for bankruptcy.

Meanwhile, the warehouse clubs continued to pursue their selective strategy, stocking a limited array of toys mainly during the Christmas season. A warehouse club executive reported to this author that, following the FTC order, his firm had no special difficulty securing the toys that it wanted. The largest warehouse club chain, Costco, expanded its number of stores from 374 in 2002 to 474 at the end of 2006. However, its principal independent rival, BJ’s, experienced disappointing sales in 2006 and a resignation by its CEO, accompanied by rumors of a possible takeover. The rise of Wal-Mart, in particular, had created a new competitive environment for both specialist and general retailers.

REFERENCES


CASE 17


INTRODUCTION

State Oil, a distributor of gasoline, entered into a three-year gasoline supply agreement with Barkat Khan in January 1992. By the terms of the supply agreement, Khan was required to purchase all of his gasoline from State Oil. The supply agreement also effectively established the maximum retail price that Khan could charge for gasoline. After Khan’s business failed, he sued State Oil, claiming that the supply contract constituted an illegal maximum vertical price-fixing agreement.

The dispute between Khan and State Oil eventually reached the Seventh Circuit Court of Appeals and the Supreme Court. The court of appeals ruled that State Oil’s contract with Khan was illegal under the appropriate Supreme Court precedent from 1968, Albrecht v. Herald Co., 390 U.S. 145. But the court of appeals strongly criticized the 1968 decision and encouraged the Supreme Court to overturn it. The Supreme Court took the opportunity presented to it by the State Oil-Khan dispute to reexamine Albrecht and agreed with the court of appeals that the earlier case had been decided incorrectly. As a result, Albrecht was overturned, and contracts between a distributor and a retailer that specify a maximum retail price are no longer always illegal.

Antitrust claims in the United States are decided under one of two standards, typically referred to as "per se" or "rule of reason." Although in practice the distinction between per se and rule-of-reason offenses often is

Gustavo E. Bamberger*

*Bamberger submitted an expert report on behalf of Khan. The author thanks John Kwoka and Lawrence White for their helpful comments.
not sharp, a per se offense refers to conduct that courts have found is almost always anticompetitive—the classic example is horizontal price fixing by members of a cartel. In the case of a per se violation, no defense is allowed for the practice, and a court does not need to investigate the competitive effects of the challenged practice.

In contrast, the rule of reason is used to investigate conduct that may or may not be anticompetitive. As the Seventh Circuit Court of Appeals has explained: “Challenged practices that do not fall within any of the per se categories are subject to the broader-ranging inquiry into effect and motives that goes by the name of the ‘rule of reason’ and that requires the plaintiff to prove that the defendant’s conduct actually (or with a high likelihood) reduced competition.”

The treatment of “vertical” restrictions under U.S. antitrust law—for example, restrictions imposed by a manufacturer or distributor on its dealers—has long been unsettled and controversial. The antitrust treatment of exclusive territories (a form of nonprice vertical restriction), for example, has varied over time. In White Motor Company v. U.S., 372 U.S. 253 (1963), the Supreme Court ruled that territorial restrictions imposed by White Motor on its distributors should be evaluated under the rule of reason. But four years later, in U.S. v. Arnold, Schwinn & Company, 388 U.S. 365 (1967), the Supreme Court ruled that exclusive territories were per se illegal—according to the Court, the use of such restrictions was “so obviously destructive of competition that their mere existence is enough.” In Continental TV v. GTE Sylvania Inc., 433 U.S. 36 (1977), the Supreme Court overruled its prior Schwinn decision and reinstated the rule of reason for exclusive territory cases.

Antitrust rulings in vertical price restriction cases have an even longer history. The Supreme Court first addressed the issue in Dr. Miles Medical Co. v. John D. Park & Sons Company, 220 U.S. 373 (1911). Dr. Miles involved the practice of “resale price maintenance,” in which a manufacturer typically specifies the minimum price that a distributor or retailer can charge for the manufacturer’s product. In Dr. Miles, the Supreme Court found that such vertical contracts were per se illegal. Congressional legislation in 1937 and 1951, however, allowed states to pass “fair trade” laws, under which resale price maintenance was allowed. These laws were repealed in 1975, and resale price maintenance has been per se illegal since then.

"Even horizontal price-fixing agreements may escape per se treatment. In U.S. v. Brown University, et al, for example, the Court of Appeals for the Third Circuit ruled that a horizontal pricing agreement between nonprofit institutions should be evaluated under the rule of reason; see the discussion in Bamberger and Carlton (2004).

Khan v. State Oil, 93 F.3d 1358 (7th Cir. 1996), at 1361.

See Carlton and Perloff (2000, at 637-639) for a brief discussion of antitrust rulings in exclusive territory and resale price maintenance cases.
The Supreme Court also addressed the use of *maximum* prices in vertical agreements. *Albrecht v. Herald Co.*, 390 U.S. 145 (1968) examined contracts between the publisher of the *St. Louis Globe-Democrat* and newspaper distributors. The *Globe-Democrat* provided its distributors with exclusive territories for the home delivery of newspapers and specified the maximum retail price that a distributor could charge. The Supreme Court ruled that the contract constituted a vertical price-fixing agreement and was thus a per se violation of Section 1 of the Sherman Act.

The Court’s *Albrecht* decision was widely criticized for years. In the 1970s, for example, Robert Bork (1993, p. 281), in his criticism of several of the Supreme Court’s rulings in vertical price restriction cases, complained that “the Supreme Court has even gone so far as to declare per se illegal a manufacturer’s attempt to fix *maximum* resale prices. . . . There could, of course, be no anticonsumer effect from such price fixing.” Similarly, Richard Posner (1976, p. 157)—who would later play a key role in *State Oil v. Khan*—commented that “[i]t is difficult to understand what proper antitrust question is raised by [the defendant’s conduct in *Albrecht]*.” Notwithstanding these criticisms, *Albrecht* remained the law of the land until the Supreme Court’s *State Oil Co. v. Khan* decision.

In the following sections, I first summarize the factual background of the dispute between State Oil and Khan. I then review the history of the case; I also discuss the antitrust treatment of vertical maximum price restrictions in the light of *State Oil v. Khan*. Finally, I revisit *State Oil v. Khan* in the context of a rule-of-reason analysis.

**BACKGROUND**

In January 1992, State Oil entered into a three-year lease with Barkat Khan for the operation of a gasoline station/convenience store in Addison, Illinois, a western suburb of Chicago located in DuPage County; at the same time, Khan agreed to be exclusively supplied with gasoline by State Oil. By the terms of the supply agreement, the wholesale price paid by Khan to State Oil was determined in relation to a recommended retail price established by State Oil:

The Landlord [State Oil] shall charge the Tenant [Khan] for each transport of gasoline at the recommended retail pump price less margin and sales tax. The Tenant’s margin will be $0.0325 (three and \( \frac{4}{100} \) cents) per gallon. In the event the Tenant elects to sell at retail at a price in excess of the recommended retail pump price established by the Landlord, the Tenant shall pay the Landlord a sum equal to the difference between the retail pump price at which the gasoline in question has been sold, and the recommended retail pump price established by the Landlord multiplied by the number of gallons sold. In the event the Tenant elects to sell at a retail
Thus, if Khan sold gasoline at a price higher than the recommended retail price, his margin would nonetheless equal 3.25 cents per gallon; if Khan set a price below the recommended retail price, his margin would be less than 3.25 cents per gallon. As a result, Khan could earn at most 3.25 cents per gallon on his gasoline sales.

In addition to the margin he earned from the sale of gasoline, Khan earned a margin on convenience store sales. In 1992, Khan’s gross margin from convenience store sales accounted for about 65 percent of the total gross margin he received from operating the gasoline station/convenience store.

Pursuant to his lease with State Oil, Khan had to make a monthly payment to State Oil. By the end of 1992, Khan was not able to meet his lease obligations. At the end of January 1993, State Oil terminated Khan’s lease and petitioned a state court to appoint a receiver to operate the former Khan location. A receiver was appointed and ran the station until September 1993.

The receiver was an experienced operator of gasoline stations who did not consider himself bound by the terms of Khan’s supply contract with State Oil. By the beginning of May 1993, he had decided to set retail prices so that he received a lower margin on sales of regular gasoline but a higher margin on sales of premium gasoline. He believed that by doing so, he would increase traffic through the gasoline station and convenience store and receive a higher average margin per gallon of gasoline sold. That is, he expected the demand for regular gasoline to be more elastic than the demand for premium gasoline. As a result, he expected to increase total gasoline and convenience store sales.

Information on the receiver’s retail price and margin per gallon by type of gasoline was not maintained. However, information was retained on gallon sales by type of gasoline and average margin per gallon. This evidence suggested that the receiver’s new pricing policy was successful. First, as expected, sales of regular gasoline as a percentage of total gasoline sales increased substantially. For example, during the five-month period May to September 1992—when Khan operated the station—regular gasoline sales

---


Khan and the receiver each sold three types of gasoline—regular, premium and “mid-grade.” Relatively little mid-grade gasoline was sold; the receiver set the same margin on mid-grade and premium. For the purpose of my discussion, I combine mid-grade and premium (into premium).

My description of the receiver’s actions are based on a discussion with the receiver.
accounted for about 55 percent of total gallons sold; one year later, during the five-month period May to September 1993—after the receiver stopped pricing so as to earn 3.25 cents per gallon of gasoline sold—regular gasoline sales were about 68 percent of the total. Second, total gasoline sales were 20.8 percent higher in May-September 1993 than in May-September 1992.  

The available evidence indicated that the 20.8 percent increase in gasoline sales was attributable, in part, to the receiver's revised pricing. According to the receiver, there were no substantial changes in competitive conditions between 1992 and 1993—for example, no gasoline stations were closed during 1993 in the immediate vicinity of the former Khan location. However, total demand in the area for gasoline apparently was higher in 1993 than in 1992. The Illinois Department of Revenue reports monthly gasoline sales by county, and that information showed that gasoline sales in DuPage County were 11.6 percent higher in May through September 1993 than in the comparable five-month period in 1992. Thus, the receiver's sales increased by a substantially higher percentage than sales in the county in which the station was located.  

Finally, the overall margin earned by the receiver on all sales of gasoline for the May through September 1993 period was 3.82 cents per gallon, or 17.5 percent higher than the maximum that Khan could have received under the terms of his supply contract with State Oil. This increase in the overall average margin shows that the receiver's increase in margin (i.e., more than 3.25 cents per gallon) for sales of premium gasoline must have been substantially higher than the receiver's reduction in margin for regular gasoline sales.

**HISTORY OF THE CASE**

**District Court**

Khan brought suit against State Oil in federal district court, claiming that the supply contract between Khan and State Oil violated Section 1 of the Sherman Antitrust Act.  Khan did not argue that the supply contract was illegal because it resulted in an anticompetitive harm, such as higher prices to consumers. Instead, Khan argued that the contract was a "clear-cut case

1 The statistics in the text are derived from financial records provided by Khan and the receiver and were reported in a damage analysis filed on behalf of Khan.
2 The former Khan location accounted for only a small percentage of total sales in DuPage County. Information for smaller areas (e.g., Addison, Illinois, the town in which the gasoline station was located) was not available.
3 The litigation at the district court and court of appeals levels was Khan v. State Oil. Because State Oil appealed the court of appeals verdict to the Supreme Court, the Supreme Court case is known as State Oil v. Khan. I refer to all of the litigation as State Oil v. Khan.
of vertical price fixing” and thus per se illegal.” In particular, Khan argued that the “recommended” retail price was effectively a maximum price because Khan would have no economic incentive to charge more than that price.

State Oil denied Khan’s claim. First, it denied that the supply contract constituted vertical price fixing because under its terms Khan was free to charge any price he chose for gasoline. Second, it argued that even if the supply contract constituted vertical price fixing, it should be evaluated under the rule of reason, not under the per se standard. Furthermore, because Khan had failed to present any evidence that the contract was anticompetitive—as would be required in a rule-of-reason analysis—Khan’s complaint should be dismissed.

The district court agreed with State Oil that the legality of the supply agreement should be evaluated under the rule of reason. Because Khan failed to provide any evidence that the supply contract had an adverse effect on competition, the district court granted State Oil’s motion for summary judgment on Khan’s price-fixing claim.

**Court of Appeals**

Khan appealed the district court’s decision to the Seventh Circuit Court of Appeals. In a decision written by Chief Judge Richard Posner—who had criticized Albrecht twenty years earlier—the court of appeals overruled the district court. The court of appeals agreed with Khan that the supply agreement between State Oil and Khan constituted a vertical price-fixing agreement and that vertical price-fixing agreements were per se illegal. The court of appeals based its conclusion that vertical maximum price-fixing was per se illegal on Albrecht. But the court of appeals also argued that Albrecht was mistaken and should be overturned.

The court of appeals began by investigating State Oil’s claim that the supply agreement did not set a maximum retail price because Khan was allowed to charge a higher price so long as he turned over any proceeds that reflected a greater than 3.25 cent per gallon margin to State Oil. The court of appeals was not convinced by the State Oil claim:

> State Oil. . . denies that the provision in the contract pertaining to Khan’s charging a price above the suggested retail price is a form of price fixing. It points out that Khan was free to charge as high a price as he wishes. This is true in the sense that it would not have been a breach of contract for Khan to raise his price. But the contract made it worthless for him to do so. . . . Generally when a seller raises his price, his volume falls; and if his profit on each unit sold is frozen, the effect of his raising...”

“Memorandum in Support of Plaintiffs’ Motion for Summary Judgment on Count II,” at 3.
price will be that he loses revenue: he will sell fewer units, at the same profit per unit.”

The court of appeals concluded that "State Oil engaged in price fixing."

The court of appeals then turned to whether maximum price fixing is illegal per se. The court of appeals noted that a distributor (e.g., State Oil) may want to impose a maximum retail price on a dealer (e.g., Khan) for pro-competitive reasons. For example, if State Oil granted Khan an exclusive territory—so as to encourage more efficient marketing—and Khan thereby gained some measure of market power, a maximum retail price could prevent Khan from exploiting that market power. State Oil would not share in any of the proceeds generated by an exercise of market power by Khan—indeed, such an exercise of market power would harm State Oil by reducing the amount of gasoline it sold to Khan. Thus, State Oil would have an incentive to constrain the exercise of market power by Khan through the use of a maximum retail price. In this case, State Oil’s incentives would be aligned with those of consumers.

But the court argued that despite the potential pro-competitive explanations for setting a maximum retail price

the Supreme Court has thus far refused to reexamine the cases in which it has held that resale price fixing is illegal per se regardless of the competitive position of the price fixer or whether the price fixed is a floor or a ceiling. The key precedent so far as the present case is concerned is Albrecht v. Herald Company. . . , a damages suit like this where the Court held over a vigorous dissent that the action of a newspaper publisher in fixing a ceiling at which its distributors could resell the newspaper to the public was illegal per se.”

Although State Oil argued that Albrecht was no longer the view of the Supreme Court, the court of appeals rejected State Oil’s position. Instead, the court of appeals concluded that the district court had erred and that Khan’s “antitrust claim should not have been dismissed.”

But the court of appeals made clear that it believed that Albrecht was bad antitrust law:

[D]espite all its infirmities, its increasingly wobbly, moth-eaten foundations. . . . Albrecht has not been expressly overruled. . . . And the Supreme Court has told the lower federal courts, in increasingly emphatic, even strident, terms, not to anticipate an overruling of a decision by the Court; we are to leave the overruling to the Court itself. . . . Albrecht was unsound

Khan v. State Oil, 93 F.3d 1358 (7th Cir. 1996), at 1360.
"Khan did not claim that he had been granted an exclusive territory by State Oil.
Khan v. State Oil, 93 F.3d 1358 (7th Cir. 1996), at 1362.
"Khan v. State Oil, 93 F.3d 1358 (7th Cir. 1996), at 1365.
when decided, and is inconsistent with later decisions by the Supreme Court. It should be overruled. Someday, we expect, it will be.’

**Supreme Court**

Not surprisingly—given the court of appeals’ unambiguous encouragement—State Oil appealed the Seventh Circuit’s decision to the Supreme Court. In its decision, the Supreme Court accepted the court of appeals’ analysis and took up its invitation to reconsider Albrecht. In a unanimous decision, the Supreme Court overruled Albrecht.

The Court’s analysis began by reviewing relevant Supreme Court precedents: “A review of this Court’s decisions leading up to and beyond Albrecht is relevant to our assessment of the continuing validity of the per se rule established in Albrecht.” The Court began with its Dr. Miles decision, in which it recognized the illegality of agreements under which manufacturers or suppliers set the minimum resale prices to be charged by their distributors.

In subsequent cases, the Court’s attention turned to arrangements through which suppliers imposed restrictions on dealers with respect to matters other than resale price. In White Motor Co. v. United States [citation omitted], the Court considered the validity of a manufacturer’s assignment of exclusive territories to its distributors and dealers. The Court determined that too little was known about the competitive impact of such vertical limitations to warrant treating them as per se unlawful. Four years later, in United States v. Arnold, Schwinn & Co. [citation omitted], the Court reconsidered the status of exclusive dealer territories and held that… a supplier’s imposition of territorial restrictions on the distributor was “so obviously destructive of competition” as to constitute a per se violation of the Sherman Act.”

The Albrecht decision came in the year after Schwinn, and the Court noted that the Albrecht Court’s ruling was influenced by prior decisions in Schwinn and related cases.

Nine years after Albrecht, the Court overruled Schwinn in GTE Sylvania. The Court explained that the GTE Sylvania decision was, in part, based on the widespread criticism of Schwinn. The Court also noted that the GTE Sylvania decision was informed by “scholarly works supporting the economic utility of vertical nonprice restraints.”

---

16 Khan v. State Oil, 93 F.3d 1358 (7th Cir. 1996), at 1363.
The Court recognized that Albrecht had been met with the same types of criticism as Schwinn and again relied on such criticisms to overturn a precedent. "Thus, our reconsideration of Albrecht's continuing validity is informed by several of our decisions, as well as a considerable body of scholarship discussing the effects of vertical restraints. . . . So informed, we find it difficult to maintain that vertically-imposed maximum prices could harm consumers or competition to the extent necessary to justify their per se invalidation."20

Thus, in Schwinn, the Supreme Court deemed a vertical restriction—the use of exclusive territories—per se illegal but then overruled its precedent, based in large part on criticism by economists and antitrust scholars. Similarly, the Court characterized maximum retail price constraints as per se illegal in Albrecht, but used State Oil v. Khan to overrule its precedent, again based in large part on economics-based criticisms of the earlier decision.

In contrast, the per se illegality of minimum retail price constraints remained in place, although that standard also had been widely criticized—indeed, the same types of arguments that were directed at Schwinn and Albrecht had been used to criticize the Supreme Court's position in minimum price constraint cases. The Khan Court made clear that it continued to distinguish between minimum and maximum vertical price restrictions and that minimum vertical price restrictions remained illegal per se.21

Ten years after its Khan decision, however, in Leegin, the Supreme Court reconsidered the per se treatment of minimum vertical price restrictions. Leegin, the manufacturer of the Brighton brand of women's accessories, instituted a retail pricing policy in 1997 in which it made clear that it would not sell to retailers that discounted Brighton products. In 2002, Leegin learned that PSKS, a specialty retailer, had placed Brighton products on sale and so suspended Brighton shipments to PSKS. PSKS brought suit, claiming that Leegin engaged in an illegal vertical price fix with retailers.22

In federal district court, a jury found against Leegin and awarded PSKS substantial damages for lost sales. Leegin appealed, claiming that the rule of reason should have been used to evaluate PSKS's antitrust claims. The Court of Appeals for the Fifth Circuit disagreed:

Leegin asserts: although the Supreme Court first applied the per se rule to vertical price fixing in Dr. Miles Medical Co. v. John D. Park & Sons Co. [citation omitted], it has not applied the rule consistently. The cases cited by Leegin in which the Court applied the rule of reason, however, did not

involve a vertical minimum price-fixing agreement. See State Oil v. Khan [citation omitted] …

Because the Court has consistently applied the per se rule to such agree­ments, we remain bound by its holding in Dr. Miles Medical Co."

Leegin appealed to the Supreme Court. In a 5–4 decision, the Court agreed with Leegin. Based on the "economics literature [which] is replete with procompetitive justifications for a manufacturer's use of resale price maintenance," the Court found that:

Vertical agreements establishing minimum resale prices can have either procompetitive or anticompetitive effects, depending upon the circum­stances in which they are formed. And although the empirical evidence on the topic is limited, it does not suggest efficient uses of the agreements are infrequent or hypothetical … As the rule would proscribe a significant amount of procompetitive conduct, these agreements appear ill suited for per se condemnation."

The Court concluded "that Dr. Miles should be overruled and that vertical restraints are to be judged by the rule of reason.""

VERTICAL MAXIMUM PRICE RESTRICTIONS
AFTER STATE OIL v. KHAN

Although the Supreme Court abandoned the per se standard for vertical maximum price restrictions, it did not go as far as clearing such restrictions of all antitrust scrutiny. In particular, it refused to adopt a standard of per se legality:

In overruling Albrecht, we of course do not hold that all vertical maximum price fixing is per se lawful. Instead, vertical maximum price fixing, like the majority of commercial arrangements subject to the antitrust laws, should be evaluated under the rule of reason. In our view, rule-of-reason analysis will effectively identify those situations in which vertical maxi­mum price fixing amounts to anticompetitive conduct."

Some commentators have criticized the Court’s decision to apply the rule of reason to maximum vertical price restrictions. Blair and Lopatka

127 S.Ct. 2705, at 2710.
(1998, p. 179), for example—whose prior writings on vertical restrictions were cited by the Khan Court—argue that

[although Khan contributed significantly to sound antitrust policy, the decision is flawed. It replaced the rule of per se illegality, not with a rule of per se legality, but with the rule of reason. In the context of maximum vertical price fixing, a test of reasonableness makes little sense. The practice has no discernable anticompetitive consequences in any setting. Further, the content of the test is a mystery, and therefore its application is unpredictable.

In what circumstances could vertical maximum price fixing be found to be anticompetitive under the rule of reason? In its Khan decision, the Court referred to the justifications given for the per se treatment of maximum vertical price restrictions in the Albrecht decision and argued that the conduct described in Albrecht "can be appropriately recognized and punished under the rule of reason." Thus, the Court suggested that the anticompetitive explanations given for vertical maximum price restrictions in Albrecht could be the basis for a rule-of-reason finding of illegality. However, none of the Supreme Court's three Albrecht explanations seems likely to support a rule-of-reason-based finding of anticompetitive effect.

The first explanation of anticompetitive effect provided by Albrecht is that "maximum prices may be fixed too low for the dealer to furnish services essential to the value which goods have for the consumer or to furnish services and conveniences which consumers desire and for which they are willing to pay." But this argument implies that the manufacturer or distributor would take actions that were not in its self-interest. Indeed, the Khan Court recognized the flaw in this theory: "such conduct, by driving away customers, would seem likely to harm manufacturers as well as dealers and consumers, making it unlikely that a supplier would set such a price as a matter of business judgment.""

The second argument advanced by Albrecht is that "maximum price fixing may channel distribution through a few large or specifically advantaged dealers who otherwise would be subject to significant nonprice competition." This argument is similar to the first and likely to be insufficient to support a finding of anticompetitive effect for the same reason—a manufacturer would not be expected to take actions that limit its dealer network, unless it were efficient to do so. Again, the Khan Court recognized the flaw in Albrecht's explanation: "It is unclear, however, that a supplier would profit from limiting its market by excluding potential dealers. . . . Further,
although vertical maximum price fixing might limit the viability of inefficient dealers, that consequence is not necessarily harmful to competition and consumers."

The final Albrecht explanation relies on converting a maximum vertical price restriction into a minimum vertical restriction. In particular, Albrecht argued that "if the actual price charged under a maximum price scheme is nearly always the fixed maximum price, which is increasingly likely as the maximum price approaches the actual cost of the dealer, the scheme tends to acquire all the attributes of an arrangement fixing minimum prices." This explanation also seems unlikely to support a finding of anticompetitive effect in a rule-of-reason evaluation of a maximum vertical price-fixing claim.

If maximum vertical restrictions are in place, a dealer is not prevented from cutting price below the maximum. Albrecht avoided this difference between maximum and minimum prices by referring to situations in which "the maximum price approaches the actual cost of the dealer." Although it is correct that a dealer would typically have no incentive to cut its price if the maximum price imposed on it equaled the dealer's costs, there is, in general, little or no reason to be concerned about minimum price restrictions if the minimum price is set at or near cost. Thus, in a situation in which a minimum price could be objectionable—that is, a minimum price substantially above dealer cost—the dealer whose maximum price is restricted could cut price, and a maximum vertical price restriction is unlikely to be equivalent to a minimum vertical price restriction.

Other explanations not discussed by the Court could, however, provide a basis for a finding that a maximum vertical price restriction could be anticompetitive. One potential antitrust attack on maximum vertical price restrictions was discussed by the court of appeals. The court of appeals notes that "[t]he difference between what a supplier charges his dealer and what the dealer charges the ultimate customer is, functionally, compensation to the dealer for performing the resale service." That is, the margin that a dealer earns can be thought of as the "price" that the dealer charges for performing the resale function. Thus, by setting a maximum retail price, a manufacturer could reduce the "price" it pays for resale service, perhaps below the otherwise market-clearing price. If the manufacturer is competing with other manufacturers for dealers, this is an unprofitable strategy. But if a manufacturer is a monopolist, such a policy could reduce output of the final good, leading to higher prices, lower distribution costs, and higher

\(^{32}\)State Oil v. Khan, 188 S. Ct. 275 (1997), at 283.


\(^{34}\)As I have discussed, there was widespread criticism of the Supreme Court's position that minimum vertical price restrictions were per se illegal.

\(^{35}\)Khan v. State Oil, 93 F.3d 1358 (7th Cir. 1996), at 1361.
profits for the manufacturer. That is, a monopolist manufacturer could exercise monopsony power over its dealers.

Although this explanation is correct as a matter of economic theory, it seems unlikely to be of practical importance in many, if not most, industries. The exercise of monopsony power would be possible only if a monopolist manufacturer faced an upward-sloping supply curve for buying dealer services. If not, the manufacturer could not reduce the price of dealer services. But the supply curve of dealer services—at least in the long run—is unlikely to be upward sloping. As Blair and Lopatka (1998, pp. 161-162) explain:

Notice that, if the buyers in question face a horizontal supply curve, that supply curve will coincide with the marginal factor cost curve, and the buyers will not be able to affect price by reducing the quantity purchased. A positively sloped supply curve implies that the resources necessary to provide the services in question are specialized to the needs of the buyers under consideration. Thus, if the buyer demanded less, the supplier would not be able to shift to an alternative buyer. The resources necessary to provide distribution services, however, tend not to be specialized to the needs of the buyers under consideration. If a single purchaser of some kind of distribution services attempted to depress price, dealers would switch to other buyers desiring the same kind of services; if a group of firms in the same industry conspired to offer a depressed price for distribution services, dealers would switch to buyers in other industries.

A different potential attack on such restrictions is provided by Blair et al. (2000, p. 221), who conclude that "one can unequivocally endorse jetisoning the Albrecht rule of per se illegality, and that the Supreme Court's decision to retain a rule of reason was correct. Allowing the pendulum to swing to per se legality as many economists seem to advocate would have been a mistake." The Blair et al. (2000) attack on maximum vertical price restrictions is based on a generalization of the "successive monopoly" model of vertical price restrictions.

A standard (pro-competitive) explanation for the use of maximum vertical price restrictions is that they can be used to solve the "successive monopoly" problem. Suppose that a manufacturer has market power and that the demand for its product can be increased by the provision of dealer services. To encourage the provision of these services, the manufacturer may choose to award exclusive territories, so as to prevent one dealer from free-riding on the promotional efforts of another dealer. But by creating exclusive territories, the manufacturer creates market power for its

"In general, the exercise of monopsony power requires an upward-sloping supply curve. Carlton and Perloff (2000, p. 187), for example, explain that "[i]long-run monopsony power is impossible if the long-run supply curve is flat, because then price cannot be lowered below the competitive price."
distributors. To prevent the exercise of market power by its dealers, the manufacturer uses maximum vertical price restrictions.

Blair et al. (2000) develop a model in which the dealer chooses the retail price and a level of service provision. They assume that the manufacturer cannot control the amount of service provided by dealers. By imposing vertical price restrictions, the manufacturer can prevent the exercise of market power by its dealers. But the dealers may respond to the reduction in retail price by reducing the provision of service—because each additional sale becomes less profitable if retail price is constrained, the dealer’s incentive to provide service is reduced.

Because a reduction in dealer services will tend to reduce consumer demand, and the reduction in retail price will tend to increase consumer demand, Blair et al. (2000, p. 217) argue that the net effect of maximum vertical price restrictions is ambiguous: “Absent service at the distributor stage, maximum resale price restraints always lead to an increase in output. When we generalize the model to include service by the distributor, however, the effect on output is ambiguous.” Blair et al. (2000) argue that because consumer welfare increases as output increases, a rule-of-reason approach to maximum vertical price constraints implies an “output test”—if output increases as a result of imposing a maximum retail price, the practice is pro-competitive; but if output declines, the imposition of maximum retail prices should be condemned under the rule of reason.

Grimes (2006, 851) proposes a different attack on vertical maximum price fixing, based on the “post-Chicago literature on distribution restraints.” Grimes criticizes the State Oil v. Khan decision—and an amicus brief filed by the United States on behalf of State Oil—for ignoring this literature, which Grimes characterizes as “demonstrat[ing] that vertical maximum price fixing is an upstream power restraint much more closely related to ties than to vertical minimum price-fixing.” However, Grimes does not claim that the Supreme Court was wrong to overturn Albrecht. Instead, he concludes that “[w]ith the benefit of [the post-Chicago] insight, the Court could have written an opinion that provided much more meaningful guidance on how vertical maximum price fixing cases should be tried under the rule of reason.”

It is not clear how courts will apply the rule of reason to maximum vertical price-fixing cases in the aftermath of Khan. In one federal district court case, for example, the court cited Khan and explained that:

At a basic level, maximum resale price maintenance may be consumer-friendly, as the practice evidences a manufacturer’s desire to keep prices below a stated point and closer to competitive levels. . . . Given these

"The Blair et al. (2000) model assumes that the upstream firm is a monopolist. It is unclear whether the authors would apply their output test in a situation where the firm that imposes the vertical restrictions competes with other upstream firms.
potentially pro-competitive consequences of maximum resale price main-
tenance, it follows that plaintiffs should be put to the task of demonstrat-
ning that the alleged restraint has identifiable anticompetitive effects in a
relevant market."

The district court did not, however, discuss what type of analysis or
evidence could establish an "identifiable anticompetitive effect." Instead,
the court found that plaintiffs had failed to define an appropriate product or
geographic market—which "makes it impossible to weigh the effects of the
[defendant's] alleged conduct"—and so dismissed the plaintiff's maximum
vertical price-fixing claim.

**STATE OIL v. KHAN REVISITED**

In addition to overruling *Albrecht*, the Supreme Court remanded *State Oil v. Khan* to the court of appeals "for consideration of whether State Oil's
maximum price-fixing violated the Sherman Act under the 'Rule of
Reason,' as distinct from the per se rule." But Khan elected not to provide
a rule-of-reason analysis, and the court of appeals considered the argument
waived. Instead, Khan argued that the price restrictions in the supply con-
tract he signed with State Oil should be considered minimum price restric-
tions and thus per se illegal. The court of appeals concluded that Khan's
argument was "plainly without merit" and thus remanded the suit to the dis-
trict court with directions to dismiss the suit.

But what would a rule-of-reason investigation have shown if one had
been conducted? It is clear that lifting the vertical price restrictions had an
economic effect—the receiver's average margin and total sales increased.
What is less clear is why State Oil would impose a policy that reduced
its total sales. Klein (1999) provides one possible explanation. Klein
explains that State Oil may receive a higher (wholesale) margin per gallon
of premium gasoline than per gallon of regular gasoline. If that were the
case, State Oil would be interested not only in the total amount of gasoline
it sold to Khan, but also in the mix of those sales between regular and pre-
mium gasoline. In particular, State Oil would want to sell a relatively high
percentage of premium gasoline and would want to prevent Khan from
charging a high price for premium. That is, "the dealer's shift in gasoline


"Khan v. State Oil, 143 F.3d 362 (7th Cir. 1998), at 363.

"Because convenience store sales were important to profitability, Khan could have reduced the
margin on regular gasoline sales even if he were constrained to a 3.25-cents-per-gallon margin on
premium sales. Presumably, he could have increased his total sales even more than the receiver.
Whether this would have been a more profitable strategy than pricing so as to receive a 3.25-cent
margin on every gallon sold is not clear from the available evidence. See Klein (1999, p. 55).
demand towards regular and away from premium hurts the gasoline sup­plier by producing a mix of premium and regular gasoline sales that is not jointly profit maximizing [i.e., does not maximize the joint profits of State Oil and Khan]" (Klein 1999, p. 56).

Evaluating the competitive effect of the vertical maximum price restrictions imposed on Khan is complicated by the fact that Khan was selling two products—regular and premium—that were affected differently by the imposition of the price restrictions. Ignoring the different demands for the two products leads to the apparently anomalous finding that lifting the vertical restrictions resulted in both an increase in price (i.e., the average dealer margin increased) and an increase in total output.

Were the State Oil maximum vertical price restrictions anticompetitive because they resulted in lower sales of regular gasoline (although higher sales of premium gasoline)? Other contractual practices not generally con­demned by the antitrust laws can harm one group of consumers and benefit a second. To take a closely related example, consider the case of territorial restrictions adopted to encourage dealer services such as the provision of information. Territorial restrictions that result in additional information to consumers also can shift a dealer’s sales mix. By increasing the amount of information available to potential consumers, the policy will presumably tend to increase the sales of products for which information is relatively more important (e.g., those that are more complicated) at the expense of products for which information is less important. If the territorial restric­tions lead to higher retail prices for all products, then the restrictions could harm consumers who buy products for which information is not important and benefit consumers who buy information-intensive products. This type of shift in sales mix can be profitable for the manufacturer if its margin on information-intensive products is higher than for other products.

A rule-of-reason analysis generally will not condemn territorial restrictions of the type in my previous example (especially if the manufac­turer imposing the restrictions is one of many competing in a market). Thus, condemning the same type of result—that is, a vertical restriction that led to higher sales of one product but lower sales of a second—under the rule of reason because it was generated by a vertical price restriction instead of a vertical nonprice restriction would create an inconsistency in the treatment of the two types of restrictions with no basis in economics.

CONCLUSION

The per se illegality of vertical restrictions has long been criticized by economists because these practices often benefit consumers. Largely as a result of such criticism, the per se illegality of exclusive territories—a ver­tical nonprice restriction—was replaced by a rule-of-reason treatment of
the practice in 1977. Almost twenty years later, the Supreme Court used State Oil v. Khan to further narrow the range of vertical practices that is per se illegal. The Supreme Court's decision to overrule a prior decision was again largely based on economic analysis of the prior ruling. It is unclear how these practices will be analyzed under the rule of reason, however, because the Supreme Court provided little guidance on how maximum vertical price restrictions could be shown to have an anticompetitive effect.

REFERENCES


PART IV
Network Issues
Antitrust has only recently begun to address the concepts of networks and network effects directly, but networks themselves—for example, telephone, rail, pipelines, electricity, broadcasting, even highways—have been around far longer. The economic issues that they raise—economies of scale, strategic uses of assets, refusals to deal, and so on—also have long histories, and they arise in many other contexts that are described in other parts of this book. In those other contexts, antitrust has addressed network issues—often described in other terms—for a long time as well.

We will first address the economics of networks and then the relevant antitrust law.

**ECONOMICS**

**Some Basic Concepts**

Technically, a network is a set of nodes connected by links. Because this definition does not convey much intuition, some diagrams and examples will surely help.

**A Simple Network**

Figure IV-1 portrays a simple "star" or "hub-and-spokes" network. The outer nodes (A, B, C, etc.) are all connected to each other through a central node S. Examples of networks that look approximately like Figure IV-1 include:

- a local telephone exchange, with the outer nodes as the users, the central node as the central switch that connects them, and the connecting links as
the telephone lines (wires) that connect the users to the switch; a local e-mail system could similarly be portrayed;

• a small airline ("hub-and-spokes") network, with the outer nodes as smaller (less central) airports, the central node as the "hub" through which passengers must travel (and usually change planes) if they want to fly from one peripheral city to another, and the connecting links (spokes) as the flights between the smaller airports and the hub; many local (metropolitan) commuter rail and bus systems have a similar structure;

• a local package or mail delivery system, with the outer nodes as pick-up and delivery customers, the central node as the sorting and routing facility, and the links as the collection and delivery routes;

• a local electricity system, with the outer nodes as electricity users, the central node as a generating station, and the links as the wires that carry electricity to users;

• a local radio, TV, or cable broadcasting system, with the outer nodes as listeners/viewers, the central node as the broadcasting facility, and the links as the over-the-air or over-the-cable broadcast transmissions.

Some important insights can be gained from even these simple examples. First, the central nodes really are central; all transactions must travel through or originate at that node. An immediate implication is that the central node can be characterized as a potential "bottleneck" for the system.

"The telephone system is frequently described as a "switched network."

Page 475
or an "essential facility": Without access to it, operation in the market is impossible.

Further, the capacity of that node is important for the system. If there is adequate capacity at the central node (or extra capacity can be added at relatively low cost), then extra users can be added to the network at relatively low cost. In essence, economies of scale for the central node and thus for the system are an important aspect of networks.¹

These economies of scale are usually a component of economic models of networks, and they do seem to be present in many real-world networks. However, short-run capacity constraints can cause diseconomies of scale, which manifest themselves through congestion, delays, and static in these systems; and whether additional capacity can be added at modest or substantially higher costs (including managerial and other "systems" costs) is an empirical question, and not a mandate of nature or the guaranteed result from an engineer’s calculations.

Second, compatibility of the nodes and links is essential in order for transactions to occur. Compatibility is achieved through the nodes’ and links’ sharing (and adhering to) common specifications or standards. Compatibility standards may be of a physical nature:

- The telephone jacks must fit the wall receptacles;
- Rail gauge (track width) and wheel spacing, tunnel dimensions and freight car dimensions, and so on, must fit with each other;
- Airport runways must be a suitable length for the airplanes that are expected to land and take off;
- Electrical plugs must fit the wall outlets; and so on.

Compatibility standards may also be a technological phenomenon:

- The telephone equipment on a system must be capable of sending and receiving the same analog or digital information in order to be able to communicate;
- The voltage and cycle requirements of electrical appliances must be compatible with the electricity flow;
- Radio and television receivers have to be compatible with the broadcast signal; and so on.

Third, there is an important distinction between the telephone, airline, and package networks and the electricity and broadcasting networks. The former networks are two-way networks: The outer nodes are all capable

¹Scale and capacity issues may also be present for the links.

Over-the-air broadcasting is the one exception among these examples, since congestion or delays are not relevant. The broadcast signal is a true public good: It is nonrival (one viewer/listener does not prevent another from viewing/listening) and nonexclusive (in the absence of scrambling, no one can be excluded from viewing/listening); see Pindyck and Rubinfeld (2001, ch. 18).
of both sending (initiating) and receiving transactions, and transactions can flow back and forth (two-way) between any two outer nodes (via the central node). By contrast, the latter networks are one-way or distributive networks: the outer nodes are capable only of receiving transactions from the central node, so the flow of transactions is solely (one-way) outward from the central node.

Fourth, so long as economies of scale are present, two-way networks enjoy positive direct "network externalities" (Rohlfs 1974, 2001): When an extra user (an additional external node) joins the network, the other users directly benefit. Thus, if extra telephone users join a local exchange network, not only do they receive benefits from the calls that they make to and receive from others on the network, but the others benefit as well (by being able to receive and make more calls); the same effects occur for the airline and package systems.7

By contrast, the one-way networks do not generate direct network effects. There is no direct gain to one electricity user if another user joins the network (and no direct gain to one viewer/listener to another's tuning in). But there may be indirect network effects that operate through the economies of scale of the central node: More electricity users on the network permit the fixed costs of the central generating facility to be spread over the larger number of users and thus may permit lower rates to all users; more viewers/listeners may permit higher quality and a wider variety of programming to be broadcast.

Fifth, network members often have to make network-related investments: in equipment or location or in learning the technology of the system. Often these investments are "sunk"; that is, they are specific to that network and are not transferable to another (somewhat different) network. The differing compatibility standards—physical and/or technological—of networks can contribute to this sunk aspect of the investment.

A Slightly More Complex Network

In Figure IV-2 we expand the conceptual coverage of the diagrams by portraying two star networks that are connected by a "trunk" link. This expanded diagram could represent:

- two separate local telephone networks that are connected (through their two respective switches, S, and S,) by a long-distance trunk line; again, a more complex e-mail system could similarly be portrayed;

7 As Liebowitz and Margolis (1994) point out, whether the other users enjoy a net benefit will depend on whether the pricing of the network services takes these externalities into account, in which case the network owner will have internalized the externality.

8 The assumption of adequate capacity—economies of scale—is crucial. If the extra user causes added congestion for the others, then the direct network effects could be negative.
two local or regional railroad freight collection and distribution systems, with S\textsubscript{A} and S\textsubscript{B} representing marshalling yards, connected by a long-distance trunk line (or an airline route system with two hubs and a connecting route);

an electrical system, with generating units A\textsubscript{1} on the left (coordinated by S\textsubscript{A}) and electricity users B\textsubscript{1} on the right (with a voltage "step-down" facility S\textsubscript{B});

a more complete cable TV model, with cable networks A\textsubscript{1} on the left and subscribers B\textsubscript{1} on the right, with S\textsubscript{A} representing the cable system’s “packaging” function, and with S\textsubscript{B} representing the system’s distribution “head end,” from which the channels are electronically distributed to subscribers.

The characteristics of the simple star network carry over to the double star network: The central nodes, S\textsubscript{A} and S\textsubscript{B}, continue to have a bottleneck position vis-à-vis any outer node that is directly connected to it in its local system and also with respect to the outer node of any other local system that wishes to transact with an outer node in that central node’s system; and economies of scale are an important feature, as are compatibility, standards, and the possibility of significant switching costs by users. Further, the first two examples involve two-way networks and direct network effects; the latter two involve one-way networks and indirect network effects.

The double star allows us to expand the network concept to other systems that have network-like characteristics. For example, consider a system that consists of personal computer (PC) application software packages (the outer nodes A\textsubscript{1}), a PC operating system (OS) (node S\textsubscript{A}), a microprocessor

Though both use the Internet, e-mail (including instant messaging) is clearly a two-way network, while the World Wide Web is primarily a one-way network.
(node $S_1$), and PCs (outer nodes $B^\wedge$). Though this is not a physical network, it has the characteristics of a network: central nodes, economies of scale, and the importance of compatibility and standards (since all of the components must be able to function with each other). Further, the network effects are indirect (as is true for one-way networks): If an extra PC user joins the system, there is no direct benefit for other users. But there are indirect benefits: As more users join the system, more software applications developers are willing to write software for this larger volume of users. So the incumbent users benefit from a greater variety of software available to them when more users join the system.

Further, the network effects are indirect (as is true for one-way networks): If an extra PC user joins the system, there is no direct benefit for other users. But there are indirect benefits: As more users join the system, more software applications developers are willing to write software for this larger volume of users. So the incumbent users benefit from a greater variety of software available to them when more users join the system. Also, users of a particular system may well find that their investments are system-specific, if the system’s standards are different from those of other systems.

In sum, this PC-related system does indeed have the important characteristics of a network. Further, hardware-software combinations, such as DVD players and DVD disks, and compact disk (CD) players and CDs, and the combination of specialized machinery and its spare parts have similar characteristics. The term virtual network is frequently used for such non-physical systems.

Implications

There are a number of important implications of these network characteristics for market structure and behavior.

First, if the direct or indirect network effects are relatively strong, competition among networks—especially networks that are incompatible with each other—is likely to be unstable. Relative size itself will provide an important advantage to a larger network, which will cause it to attract an outsized fraction of new users (and of switchers), yielding greater size and greater advantage. A "bandwagon" or "tipping" effect operates.

The result (when network effects are strong) is likely to be a "winner-take-most" structure. Smaller networks will be able to survive only if they offer distinct features that are sufficiently worthwhile to their members to compensate for the smaller network size (and thus the smaller network effects) or if their users are locked into the network because of high switching costs.

Second, for such winner-take-most networks, there is likely to be fierce rivalry at the early stages of network development. Owners of rival networks will each hope to get an early advantage and start the bandwagon process. Such efforts are likely to include low introductory pricing (which others may claim is "predatory"); exaggerated claims for future product developments; exaggeration of current performance; and advertising.
lists of testimonials and endorsements by prominent customers; prominently announced alliances and joint ventures with some other network providers; and other efforts to create the impression of a bandwagon. In essence, the process of "competition for the market" may be crucial for the determination of the eventual structure of the market (and especially the identity of its dominant firm).

An immediate policy implication of this last point, but also a dilemma, is that antitrust enforcement may have to scrutinize carefully the early behavior and maneuverings of the network rivals, in order to prevent anti-competitive practices that could yield a subsequent structure that is then difficult to change (for the reasons discussed below). But it is during the early stages of a nascent network industry when information about the industry is the most hazy, the technological possibilities are the most vague, and the risks of harm from inappropriate interventions are the greatest.

Third, the central nodes of the network—as bottlenecks—are likely to possess market power vis-à-vis the other nodes in the network. The central node may be a physical facility, such as the central switch of a telephone network. Or the centrality of the network may be found in its compatibility standard. If the standard is proprietary, as is true for Microsoft’s Windows operating system, the owner is likely to be able to exercise market power.13 On the other hand, if the standard is in the public domain (as is true for railroad gauges and telephone jack design), a bottleneck problem is unlikely to arise.14 Important determinants of the extent of the owner’s market power will be the strength of the network effects and the size of users’ switching or lock-in costs.

Fourth, entry by new networks (if they are not compatible with incumbent networks) is more difficult than in otherwise similar nonnetwork industries, because the entrant suffers from the reverse network effects of small size and customer base. In essence, it is harder for an entrant to convince customers of other networks to sample its product or service, because the switchers lose any network-effects benefits from their old network as well as incurring any switching costs. Equivalently, the entry costs and risks for an entrant to attain a network size that approximates an incumbent may be quite large.15

Fifth, fundamental innovation involving incompatibilities with incumbent networks may be more difficult in network industries, because of the

13 See the discussion of Microsoft in Case 20 by Daniel Rubinfeld in this part.
14 This may not prevent companies from attempting to create important differences for strategic purposes. AT&T, for example, sought to ensure that only its equipment would be used by arguing that equipment from other manufacturers would jeopardize the integrity of the telephone network. See Noll and Owen (1994).
15 This appears to be the case for networks such as rail and local telephone service. Interconnection among competing networks (such as rail or telephone) can ameliorate some of these problems. Also, as is true for some Internet networks, the costs for members to belong to more than one network may be relatively low.
difficulties of attracting incumbent networks’ customers to a new product. (Innovation that is compatible with a network standard and can easily be integrated with incumbent networks, however, should face fewer problems.) Of course, cars did supplant horse-drawn carriages, CDs did replace vinyl records, and DVDs have replaced videotape formats. So fundamental innovation clearly can and does take place. Nevertheless, the process must be harder because of network effects and switching costs.

Sixth, decisions about compatibility or incompatibility among competing networks may have an important strategic component, as well as a technological component. If the owner of a network believes that its network is or will become the beneficiary of a bandwagon or tipping phenomenon, that firm is likely to be reluctant to interconnect—that is, become compatible with—rival networks. This has been AOL’s strategy with respect to its Instant Messaging e-mail system, as is discussed by Faulhaber (2004). This refusal to interconnect is the network technology equivalent of refusals to deal in the context of the vertical arrangements of Part III.

Further, if the setting of a compatibility standard is jointly determined by the members of an industry (say, through an industry association), there is the risk that political/strategic considerations may cause the choice of standard to become a vehicle for raising barriers to entry or raising rivals’ costs, as incumbents choose a standard that they know will make entry more difficult or that might favor live-and-let-live incumbents over mavericks.

Finally, because of tipping and the subsequent inertia of a network market structure that has a dominant firm, the process that yields a market “winner” offers no assurance that the outcome represents the socially most efficient (say, greatest customer net satisfaction) network. Though the tipping could occur because the best network provider gets an early lead, chance events might cause another provider to get an early lead and then (because of bandwagon effects) dominate, through a process that is often described as “path dependence.” Further, even if the best provider initially does dominate, subsequent technological change might argue for a new and improved network, which another firm could better provide. But the inertia of the network structure might foil the succession by the new provider.

In sum, networks embody special features that do raise extra concerns for antitrust policy. It is to antitrust that we now turn.

ANTITRUST

The above discussion illuminates the vast array of practices in network industries that might have anticompetitive effects. Some of these practices

---

16 An exception would be if the incumbent network provider saw the existence of even a compatible innovative node or link as a threat to its core bottleneck position, as was the case in Microsoft’s hostile reaction to Netscape’s compatible browser.
are essentially the same as those that arise in the case of nonnetwork indus-
tries. In these instances the enforcement agencies and courts have dealt with
them using the same underlying principles as would apply to nonnetwork
industries, but with attention to their network setting. Mergers between
firms in network industries are analyzed using the Merger Guidelines,
for example, but the analysis additionally reflects their cost and demand
interdependencies. Evaluations of alleged anticompetitive conduct, such
as predatory, in these industries also recognize the multimarket, multidimen-
sional nature of network competition.\footnote{See the discussions of the UP-SP merger by Kwoka and White (2004), the Bell Atlantic-NYNEX
merger by Brenner (2004), and the proposed MCI-Sprint merger by Pelcovits (2004). Also, Katz
(2002) describes the DOJ’s challenge to the proposed merger of United Airlines and US Airways in
2001 and its earlier challenge to Northwest Airlines’ acquisition of a significant ownership share in
Continental Airlines in terms of the network characteristics of these airline systems.}

In other cases, however, the competitive issues surrounding network
industries are distinctive or arise with unique properties. Here we focus on
two of the most important categories of such issues: constrained access to
the network and mechanisms to induce network tipping. Each of these is
assessed against the rule-of-reason prohibitions in Section 2 of the Sherman
Act concerning monopolization and attempts to monopolize. We will pro-
vide a brief review of the manner in which the enforcement agencies and
courts have brought this statutory language to bear on them.

The issue of access was first confronted in the Terminal Railroad case
of 1912.\footnote{See the discussion of the American Airlines predation dispute by Edlin and Farrell (2004) and by
Katz (2002).} One company controlled a key bridge as well as approaches and
terminals serving that bridge at the St. Louis crossing of the Mississippi
River. When that company was acquired by a group of the railroads using
those facilities, the now jointly owned company could and did deny access
to rival railroads that also sought to use them. The Supreme Court ulti-
mately mandated access to the bridge, terminals, and approaches to those
rivals. While this opinion is generally cited as the origins of the essential
facility doctrine, it should be noted that this ruling was motivated more by
the plight of horizontal competitors—the rival railroads that had to compete
with the vertically integrated bridge/railroad company—than simply by the
facility’s being essential. As will be discussed below, in cases where a com-
pany owns a crucial facility without also competing with buyers of its ser-
vices, the courts have adopted a different view.

Other cases involving “essential facilities”—sometimes couched in
terms of denial of access, discriminatory access, and refusals to deal with
unintegrated rivals—have involved the Associated Press’s news reports,\footnote{U.S. v. Terminal Railroad Association of St. Louis, 224 U.S. 383 (1912).} electricity sold to municipal systems by Otter Tail Power,\footnote{Associated Press v. U.S., 326 U.S. 1 (1945).} the complicated

\footnote{Otter Tail Power Co. v. U.S., 410 U.S. 366 (1973).}
pricing and payment structure for ski lift tickets in Aspen,\footnote{Aspen Skiing Co. v. Aspen Highland Skiing Co., 472 U.S. 585 (1985).} terms of access to railroad trackage,\footnote{Delaware & Hudson Railway Co. v. Consolidated Rail Corp., 902 F.2d 174.} and computer reservation systems maintained by the major airlines (Guerin-Calvert 1994). In general, where the courts have viewed access to be a true bottleneck, as opposed simply to be an impediment, they have been inclined to require access on some terms, though not necessarily equal terms. This inclination is diminished insofar as the crucial asset is the result of creative effort (e.g., as embodied in a patent), for which the courts have sought to ensure adequate returns, as opposed to the mere purchase of the asset (as in \textit{Terminal Railroad}).

The case that most clearly and specifically articulated the essential facilities doctrine was the private monopolization case brought by MCI against AT&T.\footnote{MCI Communications Corp. v. AT&T, 798 F.2d 1081 (7th Cir. 1983).} The facts are well known: In order to provide long-distance telephone service, MCI needed to originate and terminate calls on local exchanges that were for the most part owned by AT&T, with whom MCI was competing in the long-distance market. Many of MCFs early years were spent battling in the courts and regulatory agencies for rights to interconnect with AT&T's local exchanges, and the courts ultimately found in MCI's favor. The opinion set out the following criteria for a finding of an antitrust violation under the essential facilities doctrine: (1) control of an essential facility by a monopolist, (2) a competitor's "inability practically or reasonably to duplicate it," (3) denial of use of the facility to a competitor, and (4) feasibility of providing the facility. The government's contemporaneous antitrust suit against AT&T was based on similar allegations of AT&T's anticompetitive denial of access to MCI (Noll and Owen 1994). This case was also resolved on terms that made clear the basic merits of this allegation.

Where outright denial of access or discriminatory terms of access are not employed, other stratagems are sometimes pursued. Foremost among these have been so-called price squeezes, a practice designed nominally to permit rivals to acquire the critical input but in reality to do so only in a manner that undermines their ability to compete. In a classic price squeeze the integrated producer sets a price for sale of the crucial input to its final-product competitor(s) that, combined with the final product price, leaves too small a margin for viable operation of its rivals.

A prominent recent example in a network context is the \textit{Town of Concord v. Boston Edison} case.\footnote{Town of Concord v. Boston Edison, 915 F.2d 17 (1st Cir. 1990).} This case focused on actions by Boston Edison to secure wholesale price increases from state regulators, price increases that were then charged to the municipal electric companies in several towns, including Concord, Massachusetts, supplied by Boston Edison.
Combined with regulated retail prices, the higher wholesale prices left little or no operating margin to the municipal utilities. The town’s contention that it was being subjected to a price squeeze was rejected by the court. Its reasoning—that municipal electric companies are profit-maximizes, that regulation was effective, and so on—has been disputed elsewhere."

As noted earlier, most essential facility and related cases involve the relationship of a vertically integrated company to its unintegrated rivals. Where such integration is not an issue, however, the courts have permitted greater discretion to the owner of a critical asset in determining who should have access to it. Many such cases arise as competitors seek membership in an organization that provides a distinctive aspect to a product or service in the market. Prominent cases have involved access and membership in credit card networks, ATMs, and real estate listing services. A prominent example of this group of cases is the Antitrust Division’s suit against Visa and MasterCard, discussed by Robert Pindyck in Case 19 in this part.

Apart from access issues, certain practices in network industries may have a greater anticompetitive potential than in other settings. A key reason for this special concern is that networks are subject to “tipping,” so that despite early competition there may ultimately be only one surviving network, or at least one very dominant network. Anticipating this, any single one of multiple competing networks at the outset may decide to engage in practices intended to undermine its rivals and thereby propel itself into the winner’s circle. Similarly, when an incumbent dominant network is faced with the prospect of a new entrant, it may take actions designed to make such entry more difficult. Some of these practices may therefore have effects that are qualitatively or quantitatively different from nonnetwork settings.

Consider, for example, the practice of exclusive dealing, which may prevent encroachment on an incumbent’s position in any market and thereby require a weighing of its beneficial and adverse competitive effects. In the context of a network industry, the ability to prevent or deter customer switching may keep new rival networks and technologies from gaining the customer base necessary for scale economies and long-term viability. Such “foreclosure” has been at the center of cases involving competition between Nintendo and Atari and efforts by FTD to deter entry by rival florist networks.27 These actions not only distort customer choice in the short run, but also may alter the long-run makeup of the industry.

Similarly, strategic use of standards and choices with respect to compatibility can alter the long-run structure of a network industry by inducing or preventing tipping of the market in favor of a particular technology or

26 For such discussion, see Kwoka (1992). It should be noted that price squeezes are by no means limited to network industries. These have been commonly alleged as alternative methods of eliminating competition in industries ranging from petroleum to aluminum. See, for example, U.S. v. Aluminum Company of America, 148 F.2d 416 (1945).

27 For discussion of these and other cases, see Shapiro (1999).
firm. Among the prominent antitrust cases that have confronted the question of the legitimate versus possibly anticompetitive use of compatibility in network industries have been the government’s cases against IBM, AT&T, and Microsoft. For example, the government alleged that IBM sought unnecessarily to manipulate the interface between its central processing units and certain peripheral equipment in order to thwart rivals from gaining a foothold in the latter. As previously mentioned, in the AT&T case the company sought to prevent attachment of non-AT&T equipment to the network, and Microsoft has been found to have employed a variety of strategies to reduce artificially the interoperability of its rivals’ software. Among these are fee structures, product preannouncements, and long-term contracts with buyers that the courts have found can be (and, in the case of Microsoft, were) intended to deter migration of users to alternatives to its dominant Windows operating system.

When confronted with such practices, the courts find themselves in the uncomfortable role of having to render judgments about the technical necessity versus strategic motives of individual company decisions, but it seems difficult to envision how else such matters can be resolved. Of course, antitrust has had a long history of having to deal with difficult questions. In that respect network concepts and analysis are but the latest challenge that the agencies and courts have had to confront. As more satisfactory answers to some questions are provided by economics, it seems highly likely that here, as elsewhere before in antitrust, those advances will find their way into enforcement and judicial decisions.

REFERENCES


CASE 18


Dennis W. Carlton and
Hal Sider*

INTRODUCTION

Regulation and antitrust have the common aim of promoting competition and maximizing social welfare but differ widely in how this aim is implemented. Antitrust operates through general rules established by courts under the principles set forth in antitrust laws, primarily the Sherman Act, the Clayton Act, and Section 5 of the FTC Act. Regulation, on the other hand, often requires firms to comply with detailed industry-specific rules. Antitrust law attempts to deter certain conduct that allows a firm to create or preserve market power but does not treat the mere possession of market power as an offense. However, regulation, among its other aims, often has affirmative, legislatively established goals of creating competition and eliminating or constraining the market power of regulated firms.

The Supreme Court directly addressed the overlapping and sometimes conflicting goals of antitrust and regulation in its 2004 decision in Verizon Communications Inc. v. Law Offices of Curtis V. Trinko, LLP. The case involved claims that Verizon, the incumbent provider of local telephone service in New York, harmed competition and violated the Sherman Act by discriminating against AT&T, a rival telecommunications supplier that provided service by leasing portions of the Verizon network. The case was brought by the law office of Curtis Trinko (Trinko), a New York firm that

*Carlton and Sider have acted as consultants for Verizon and other local phone companies on a variety of matters including the Trinko case. The views expressed in this paper do not necessarily reflect those of the U.S. Department of Justice. We thank Thomas O. Barnett, Allan Shampine, and the editors for helpful comments.

purchased local telephone service from AT&T. In a unanimous decision, the Court concluded that the antitrust laws did not obligate Verizon to deal with AT&T and other rivals. The Court concluded that Trinko’s claim of discrimination, even if true, was not a sufficient basis to bring an antitrust claim under Section 2 of the Sherman Act. That is, the Court concluded that even if Verizon violated its regulatory obligations by failing to provide adequate assistance to rivals, these actions would not violate the antitrust laws.

Trinko narrowed the circumstances in which antitrust laws create an obligation for a firm to deal with its rivals and more clearly distinguished a firm’s regulatory obligation to deal with rivals from any such obligation that might arise under the antitrust laws. The Court relied heavily on economic reasoning in explaining its conclusions. The decision recognized that imposition of a general obligation to deal with rivals can harm social welfare by reducing a firm’s incentives to invest, innovate, and compete. It also recognized that judges are typically at a significant informational disadvantage relative to industry regulators in setting the terms of trade that arise from an imposition on a monopolist of a duty to deal with rival firms. In drawing these conclusions, the Court assessed the comparative advantage of regulation relative to that of antitrust in the formation of competition policy.

This chapter reviews the facts and circumstances of Trinko and evaluates the Supreme Court’s decision from an economic perspective. The next section of this chapter briefly addresses the competing and complementary roles of regulation and antitrust in the formation of competition policy and provides background on the regulatory obligations faced by Verizon that were at issue in the Trinko case. In the following section, we review the allegations made by Trinko and the procedural history of the case. We then review the Court’s decision from an economic perspective, highlighting key aspects of the decision and evaluating the Court’s economic reasoning. Finally, we attempt to compare from an economic perspective the Court’s ruling in Trinko with the economic logic of related cases, including the Department of Justice’s antitrust suit that led to the breakup of AT&T in 1984 and the Court’s decision in Aspen Ski, in which the Court affirmed a firm’s obligation to deal with a rival under certain circumstances.

ANTITRUST AND REGULATION UNDER THE TELECOMMUNICATIONS ACT OF 1996

The Competing and Complementary Roles of Regulation and Antitrust

Regulation and antitrust laws play both competing and complementary roles in the formation of competition policy. Regulators can attempt to

Some of the issues discussed in this section are presented in greater length in Carlton and Picker (2006).
protect consumers from situations in which competition may not work well, such as natural monopoly, and may attempt to promote competition in these industries by establishing industry-specific obligations on regulated firms as to how to deal with rivals as well as setting penalties for firms that fail to comply with these requirements.

While regulators in industries such as telecommunications, electric power, and natural gas often require regulated firms to interconnect with rivals, antitrust law has never created a universal obligation for firms to deal with rivals. For example, and as discussed below, the Supreme Court has never recognized claims under the "essential facilities" doctrine, although some lower courts have done so. The caution used by courts in creating obligations for firms to deal with rivals is consistent with the traditional goals of antitrust law of deterring conduct that creates or preserves market power (as opposed to penalizing firms that possess market power) as well as the recognition that the imposition of a duty to deal with rivals can diminish pro-competitive rivalry among firms. In contrast, regulators can and do obligate regulated firms to deal with rivals and specify the prices and other terms of trade governing these transactions.

A regulated firm's failure to satisfy its regulation-mandated obligation to interconnect with rivals can raise price and restrict output relative to the level expected in the absence of this failure. This is the hallmark of conduct that harms competition, which is the economic prerequisite for creating antitrust liability. However, imposition of a duty to deal creates a tension between the goal of discouraging actions that restrict output and raise price and the goals of preserving interfirm rivalry. In addition, the treble-damage penalties assessed under antitrust law can vastly exceed those typically imposed by regulators and raise basic questions about the specification of economically appropriate penalties for various types of regulatory and antitrust offenses.

The *Trinko* case addressed a variety of fundamental issues that relates to the economically appropriate scope of regulation and antitrust enforcement. The Court found that this scope depends upon the comparative advantages of each institution in identifying anticompetitive conduct and designing appropriate remedies.

Regulators often have specialized industry knowledge and explicitly attempt to set rules that balance the interests of regulated firms with the interests of different groups of consumers. In telecommunications, for example, regulators attempt to set prices that protect consumers and preserve incumbent firms' incentives to invest. Regulators also attempt to balance the interests of different groups of consumers by establishing different prices to different groups of consumers (e.g., specifying different prices for business and residential consumers as well as different prices in urban and rural areas). Of particular relevance to *Trinko*, the Telecommunications Act of 1996, which is enforced by the Federal Communications Commission (FCC) and state public utility commissions, requires incumbent local
exchange carriers (ILECs) to interconnect with rival firms and to lease elements of their networks to rivals and not to discriminate against rivals in providing such interconnections. State regulators require ILECs to develop systems that enable regulators to monitor their performance in providing service to competitive local exchange carriers (CLECs) in a nondiscrimi­natory manner. These systems provide the basis for regulators’ ongoing mon­itoring activities and regulatory assessments of financial penalties for ILEC failure to comply.

Even with detailed industry expertise, regulators are not always success­ful in establishing policies that properly balance the interests of consumers and firms. Regulators can err in attempting to implement technical regulatory goals, and there is a well-recognized concern that regulators will cater to special interests such as the regulated firm or specific consumer groups.

For example, attempts by state regulators to implement the FCC-mandated framework for pricing network elements led to wide differences in prices for the same network elements in different states. The data from the National Regulatory Research Institute (2003) indicate that state regu­lators in Illinois established a monthly lease rate of $2.59 per month for "local loops" (the copper wires used to connect customers to the phone net­work) in urban areas in 2003. The national average rate at that time was $10.92. In addition, many state regulators have substantially changed their rates for local loops in a relatively short span of time, although the underly­ing FCC framework being applied did not change. Between 2001 and 2003, for example, Arkansas dropped its monthly rate for local loops in urban areas from $71.05 to $23.34, and Washington, D.C., reduced its monthly rates from $10.81 to $4.29.

Since it is unlikely that these changes over time or differences across urban areas reflect cost differences alone, the data indicate that regulators have substantial latitude in mandating the terms under which firms are obligated to interconnect and suggest that regulated prices may differ sub­stantially from the economically appropriate level. This in turn can have important consequences with respect to incumbent firms’ incentives to expand and maintain their infrastructure and to invest in developing new technologies because incumbent firms know that they may be forced to share this infrastructure and investments with rivals.

In contrast to the detailed rules promulgated by regulators (with vary­ing levels of success), antitrust courts operate by attempting to establish gen­eral rules. While judicial opinions impose obligations only on the parties involved in a dispute, they often have far-reaching implications by providing guidance for firms in related situations. Administration by federal judges is both a major strength and weakness of antitrust. Compared with regulators, federal judges are further removed from pressures from special interests but, at the same time, typically do not have the industry-specific expertise required to design and implement industry-specific remedies. As discussed below, attempts by courts to administer the Modified Final Judgment (MFJ)
that settled the Department of Justice’s 1974 antitrust suit against AT&T pro-
vide a stark reminder of the difficulties faced by courts in attempting to
implement industry-specific remedies to competitive problems.

While courts and regulators each have comparative advantages and
disadvantages in implementing competition policy, judges are likely to be
poorly suited to define the terms of trade when creating an obligation for a
firm to deal with its rivals. The establishment of the terms of trade is partic-
ularly complicated in the telecommunications industry, which requires
coordination and the establishment of standards among a large number of
parties.

Verizon’s Obligations under the 1996 Telecom Act

There is a long history of antitrust scrutiny of the telecommunications
industry, including the U.S. Justice Department’s antitrust suit against AT&T,
which was settled in 1984 and is discussed below. In that case, the DOJ
claimed that AT&T, then the incumbent provider of local telephone serv-
ices, harmed competition by discriminating against rival long-distance car-
rriers, including MCI. The case was settled in 1982 with a consent decree
that, with some modification by Judge Harold Greene, became the MFJ. The
MFJ divided the old AT&T into seven large, regional Bell operating
companies (RBOCs) and the “new” AT&T, which provided long-
distance and other less-regulated telecommunications services and also produced
telecommunication equipment. Subsequently, two RBOCs—Bell Atlantic
and Nynex—merged and, together with GTE, formed Verizon (which later
bought MCI). Three other RBOCs—SBC, Pacific Telesis, and Ameritech—
also merged in a series of transactions and later acquired the “new” AT&T, retaining the AT&T name. The new AT&T has since merged with
BellSouth, another RBOC. The remaining RBOC, US West, merged with
Qwest, and the combined company retained the Qwest name.

As discussed in more detail below, the MFJ was overseen by the courts
from 1984 until Congress enacted the Telecommunications Act of 1996.
The 1996 Act attempted to create competition in the provision of local tele-
phone services by establishing a “carrot and stick” approach that linked
entry by RBOCs into the provision of long-distance service to their success
in opening their local networks to competition. In order to meet these goals,
ILECs were required to meet a number of conditions established by the
FCC and state regulators relating to leasing portions of their network and
providing interconnection to competitive local exchange carriers. These
conditions include factors such as complying with the network-sharing
obligation of the 1996 Act, providing rivals with "nondiscriminatory"

See also Noll and Owen (1994).

access to the incumbent’s network, and providing regulators and rivals access to operations support systems (OSS). These systems are used by ILECs to, among other things, monitor performance in meeting service requests by affiliated and nonaffiliated entities. Both the FCC and state public utility commissions created financial penalties for firms that failed to meet regulatory requirements for opening their networks to competition. CLECs and ILECs entered into a wide variety of interconnection agreements governing the mechanics of the operational aspects of their dealings.

The 1996 Act also established a general framework for establishing the price at which ILECs were to lease network elements to CLECs, which was to be implemented by state regulators. The establishment of the appropriate price for network elements generated widespread debate among economists and further generated substantial litigation. The FCC’s framework for pricing leased network elements was established “to give aspiring competitors every possible incentive to enter local retail telephone markets.” Thus, the 1996 Act required that ILECs such as Verizon deal with rivals and also required that ILECs lease network elements to these rivals at rates well below those that would be established through arms-length negotiation. The data on the regulated prices of one set of network elements—local loops—summarized above show massive price differences across states, suggesting that regulator-established rates created sizable subsidies for entrants in at least some states.

Under these circumstances, an ILEC’s failure to comply with the requirements to provide inputs to rivals at these rates would be expected to raise price and reduce output relative to the level that would have existed if the ILEC complied fully with these requirements. However, if rates are not set at economically appropriate levels, an ILEC’s failure to comply with regulatory requirements would not necessarily have an adverse effect on consumer welfare relative to that expected based on economically appropriate interconnection prices.

Finally, the 1996 Act also included an explicit antitrust “savings clause” that established that the 1996 Act did not preempt antitrust laws. The Act stated: "[n]othing in this Act or the amendments made by this Act... shall be construed to modify, impair, or supersede the applicability of any of the antitrust laws."

THE TRINKO CASE

As noted above, Trinko is a New York City law firm that obtained local telephone services from AT&T, a CLEC providing service using certain

---

*Verizon v. FCC, 535 U.S. 467, 528 (2002).*

*47 U.S.C. Section 152, Historical and Statutory Notes.*
network elements leased from Bell Atlantic (later renamed Verizon). In 2000, Trinko filed a class action complaint in the Southern District of New York in which it claimed that Bell Atlantic:

[fullfilled orders of other Local Phone Service providers’ customers after fulfilling those for its own Local Phone Service, has failed to fill a substantial number of orders for other Local Phone Service providers’ customers substantially identical in circumstances to its own Local Phone Service Customers for whom it has filled orders, and has systematically failed to inform other Local Phone Service providers of the status of their orders with Bell Atlantic concerning [the Local Phone Service providers’] customers.]

Trinko claimed that these discriminatory actions violated both the 1996 Telecommunications Act and Section 2 of the Sherman Act and that these actions harmed competition by slowing the expansion of CLEC-provided services. The Trinko complaint was filed one day after the FCC entered into a consent decree with Bell Atlantic relating to provisioning problems similar to those identified in the Trinko complaint. As part of the consent decree, Bell Atlantic paid a $3 million fine and also agreed to pay $10 million to CLEC customers for claims related to problems in handling their orders.”

As noted above, the amended Trinko complaint alleged that members of the proposed class were damaged by Bell Atlantic’s provision of a service that was below that accorded Bell Atlantic customers. In response, Bell Atlantic argued that Trinko’s claim should be dismissed because, among other reasons, the complaint failed to meet the accepted legal standard for establishing an antitrust claim under Section 2 of the Sherman Act. As the circuit court explained, “[generally, a plaintiff can establish that a defendant violates section 2 of the Sherman Act by proving two elements. ‘(1) the possession of monopoly power in the relevant market; and (2) the willful acquisition of maintenance of that power, as distinguished from growth or development as a consequence of a superior product, business acumen or historic accident.’” Bell Atlantic argued that Trinko failed to meet the second element of this test.

Trinko also raised issues regarding whether the plaintiff, who was not a direct purchaser of service from Verizon, had legal standing to bring an antitrust claim in federal court. While this issue was not addressed in the Court’s opinion, a concurring opinion by Justice Stevens, together with Justices Souter and Thomas, concluded that Trinko did not have standing to bring a mandatory dealing claim because “...it remains the case that whatever antitrust injury respondent suffered because of Verizon's conduct was purely derivative of the injury that AT&T suffered. And for that reason, respondent's suit [...] runs the risk of duplicative recoveries and the danger of complex apportionment of damages.”

“Trinko v. Bell Atlantic Corp., 540 U.S. 398, 417. This issue is addressed in Picker (2006).”

“Amended Complaint, Trinko, v. Bell Atlantic, No. 00-1910 (SDNY) 121, 169-170, 155.”

“540 U.S. 398, 404 (2004).”

In a December 2000 opinion, the District Court for the Southern District of New York agreed with Bell Atlantic and dismissed the case. The district court concluded:

The complaint points to only one act or series of acts taken by Bell Atlantic to maintain its monopoly power: Bell Atlantic’s failure to cooperate with local competitors as required by [the 1996 Act]. Even a monopolist, however, has no general duty under the antitrust laws to cooperate with competitors. [...] The affirmative duties imposed by the Telecommunications Act are not coterminous with the duty of a monopolist to refrain from exclusionary practices.”

The U.S. Court of Appeals for the Second Circuit reinstated the antitrust claim in 2002, concluding that the district court adopted an overly narrow interpretation of Trinko’s antitrust claims. The Second Circuit concluded:

The allegations in the amended complaint describe conduct that may support an antitrust claim under a number of theories. [...] First, the amended complaint may state a claim under the ‘essential facilities’ doctrine. The plaintiff alleges that access to the local loop is essential to competing in the local phone service market, and that creating independent facilities would be prohibitively expensive. The defendant allegedly has failed to provide its competitor, AT&T, reasonable access to these facilities. [...] Second, the plaintiff may have a monopoly leveraging claim. [...] The amended complaint alleges that the defendant has monopoly power over a wholesale market in which it sells access to the local loop to telecommunications carriers. It also alleges that the defendant used that power to gain a competitive advantage in a retail market in which telecommunications carriers sell local phone service to consumers.”

The Second Circuit’s opinion was inconsistent with an earlier decision by the Seventh Circuit in Goldwasser v. Ameritech in which the plaintiff also argued that discrimination by the incumbent local exchange carrier violated Section 2 of the Sherman Act.”

305 F.3d 89, 108.
"In that case, the Seventh Circuit concluded that the alleged discrimination “cannot survive as a pure antitrust suit [...] freed from the specific regulatory requirements Congress intended. Only if Section 2 somehow incorporates the more particularized statutory duties the 1996 Act has imposed on ILECs would Ameritech’s alleged failure to comply with the 1996 Act be, in itself, also an antitrust violation.” Ibid, at 396.
The Supreme Court granted certiorari and heard arguments in the *Trinko* case in October 2003. On January 13, 2004, the Court issued a unanimous opinion concluding that "Verizon's alleged insufficient assistance in the provision of service to rivals is not a recognized antitrust claim . . . ." The Court added that "we do not believe that traditional antitrust principles justify adding the present case to the few existing exceptions from the proposition that there is no duty to aid competitors."

The Court's opinion relied extensively on economic reasoning. The Court recognized the tension between the antitrust laws' goals of fostering competition and the imposition of obligations on firms to deal with rivals. The Court noted that the desire to charge "monopoly prices" motivates pro-competitive business activity and that imposition on a monopolist of an obligation to deal with its rivals can reduce a firm's incentives to invest, innovate, and compete. The Court concluded:

> The opportunity to charge monopoly prices—at least for a short period—is what attracts 'business acumen' in the first place; it induces risk taking that produces innovation and economic growth.

> Firms may acquire monopoly power by establishing an infrastructure that renders them uniquely suited to serve their customers. Compelling such firms to share the source of their advantage is in some tension with the underlying purpose of the antitrust law, since it may lessen the incentive for the monopoly, the rival, or both to invest in those economically beneficial facilities.

As this suggests, much of the Supreme Court's reasoning, as well as its conclusion that with few exceptions "there is no duty to aid competitors," has implications beyond regulated industries. Nonetheless, the Court also stressed that the regulated nature of Verizon's activity limits the need to impose an antitrust obligation to deal with rivals.

The Court recognized that the 1996 Telecommunication Act's aim of promoting competition goes beyond the traditional goals of antitrust law. The Court highlighted the distinction between legislative and regulatory goals of promoting entry from antitrust goals of preserving competition.

> The mere possession of monopoly power, and the concomitant charging of monopoly prices, is not only not unlawful; it is an important element of the free market system. [. . .] The 1996 Act is in an important respect much more ambitious than the antitrust laws. It attempts to eliminate the

---

17 Ibid. at 407.
18 Ibid.
monopolies enjoyed by the inheritors of AT&T's local franchises. Section 2 of the Sherman Act, by contrast seeks merely to prevent unlawful monopolization. It would be a serious mistake to conflate the two goals.

The Court also stressed the regulatory nature of Bell Atlantic's interconnection requirements in choosing (again) not to recognize an "essential facilities" doctrine as the basis for creating an obligation for a firm to deal with a rival. As noted above, Trinko argued that access to Bell Atlantic's network was "essential" and that Bell Atlantic's failure to meet its regulatory obligation to provide equal access to its network violated the Sherman Act by raising price and reducing output (relative to the levels that would have existed in the absence of the violation). The Supreme Court chose not to review under the "essential facilities" doctrine accepted by some lower courts and concluded instead that Bell Atlantic's regulatory obligations undercut any need for courts to impose "forced access" remedies.

We have never recognized [an essential facilities] doctrine, and we find no need either to recognize it or to repudiate it here. It suffices for present purposes to note that the indispensable requirement for invoking the doctrine is the unavailability of access to the "essential facilities"; where such access exists, the doctrine serves no purpose [...] Respondent believes that the existence of sharing duties under the 1996 Act supports its case. We think the opposite: The 1996 Act's extensive provision for access makes it unnecessary to impose a judicial doctrine of forced access.

The Supreme Court further acknowledged that regulation is often an adequate substitute for antitrust because "[w]here such a [regulatory] structure exists, the additional benefit to competition provided by antitrust enforcement will tend to be small, and it will be less plausible that antitrust laws contemplate such additional scrutiny."

Finally, the Supreme Court also stressed that judges are at a comparative disadvantage relative to industry regulators in setting prices and other terms of trade as typically is required when a court imposes a duty to deal with a rival. The Court noted that the "incessant, complex, and constantly changing interaction of [...] would surely be a daunting task for a generalist antitrust court" and that "[a]n antitrust court is unlikely to be an effective

"Ibid.
"Complaint, Trinko v. Bell Atlantic Corp., No. 00-1910 (SDNY) 12.
"540 U.S. 398,411.
"Ibid. at 412.
"Ibid. at 414.
day-to-day enforcer of these detailed sharing obligations.” More pointedly, the Court concluded:

Enforced sharing also requires antitrust courts to act as central planners, identifying proper price, quantity and other terms of dealing—a role for which they are ill suited. Moreover, compelling negotiation between competitors may facilitate the supreme evil of antitrust: collusion.”

ECONOMIC EVALUATION OF THE COURT’S OPINION

The Supreme Court’s opinion in *Trinko* relies extensively on economic logic and arguments, both in its evaluation of the competitive effects of forced sharing rules and its evaluation of the appropriate role of antitrust in regulated industries. This section evaluates the Court’s economic reasoning.

Effect of Forced Sharing on Incentives to Invest and Compete

As noted above, there is a basic tension between the antitrust laws’ goal of encouraging competition and the goals of facilitating competition by obligating a firm to deal with rivals. As the Court recognized, this type of an obligation can discourage investment and innovation both by the monopolist, which is forced to share the results of its investments with rivals, but also by rivals. That is, if regulators set interconnection fees that do not reflect all relevant economic costs of production, then entrants have an incentive to lease the incumbent monopolists’ facilities instead of undertaking to construct their own facilities.

The effect of forced sharing rules on investment incentives has been a long-standing concern of telecommunications policy.” Requirements that

*Ibid. at 415.*

*Ibid. at 407.*

“See, for example, FCC 96-325, First Report and Order in the Matter of Implementation of the Local Competition Provisions in the Telecommunications Act of 1996, August 8, 1996 (1282), which stated, “[w]e acknowledge that prohibiting incumbents from refusing access to proprietary elements could reduce their incentives to offer innovative services.” Similarly, in congressional testimony, FCC Chairman Kevin J. Martin stated that open access regulation in telecommunications “... calls for a delicate balance: we need to make sure that incumbent networks are open to competition, but, at the same time, provide incentives for both incumbents and new entrants to build new facilities.” Kevin J. Martin, “Statement on Competition Issues in the Telecommunications Industry,” before the Committee on Commerce, Science and Transportation, United States Senate, January 14, 2003.
ILECs offer to lease unbundled network elements in turn require the specification of lease rates in addition to specifying measures of service quality. As discussed above, rates have often been set by regulators in order to subsidize competition. The obligation by ILECs to facilitate entry, and in many cases to subsidize entrants, necessarily reduces their incentive to invest in their networks and to innovate.

While the analysis of investment incentives in *Trinko* is specific to the telecommunications industry, the same general logic and economic concerns arise any time that a court requires a firm to deal with rivals. The court must take on the role of establishing not only the price at which the parties transact but also other terms of trade. In the absence of a court requirement, firms have an incentive to enter into efficient bargains if there are gains from trade. Imposition by a court of a duty to deal with a rival is required only if one party finds the deal that could be achieved through private negotiation is not as good as the court-imposed deal. Price and other contract terms established by a court will typically be at terms less favorable to the incumbent firm than those that would have been achieved without court intervention. The obligation to offer services to rivals at subsidized rates reduces a firm’s incentive to undertake the investment required to offer services in the first place.

Rivalry between incumbents and entrants also increases incentives to innovate as firms attempt to retain customers by developing and offering higher-quality products. Indeed, it can be argued that the breakup of AT&T helped spur the major technological changes in the telecommunications industry in recent decades, including the development of wireless telephony, high-speed data networks, and the Internet. Of course, it also can be argued that line-of-business restrictions in the MFJ slowed the adoption of new technologies by limiting the ability of RBOCs to offer these services. While of great interest, this question is beyond the scope of this chapter.

Nonetheless, it is important to recognize that although forced sharing requirements can help subsidize entry and innovation, forced sharing requirements can also discourage investment and innovation by the entrants. CLECs that obtain access to network elements at subsidized rates are likely to defer investments in new equipment in favor of leasing network elements from incumbent carriers. The CLEC competition that results reflects a form of "regulatory arbitrage" resulting from the establishment of economically inappropriate prices for network elements and can displace facilities-based competition. The latter would yield greater opportunities for innovation and real price competition, as opposed to price competition resulting from regulator-imposed transfer prices that subsidize the purchase of key inputs.

"Carlton and Perloff (2005), pp. 560-564.
Economic Literature on Exclusionary Conduct

In recent years, economic analysis has identified a number of circumstances in which exclusionary conduct can adversely affect competition and consumer welfare. This literature identifies a number of circumstances in which a monopolist's refusal to deal can harm competition and other circumstances in which competition would not be adversely affected. In particular, exclusionary practices by a monopolist may have the effect of preventing rival producers of complementary products from operating at an efficient scale. In addition, a monopolist may have the incentive to use exclusionary practices to slow the development of complementary new technologies that may present a future challenge to the monopolist. The IBM and Microsoft cases each raise these types of issues.

While economic understanding has grown in recent years with respect to how various business strategies, including types of exclusionary conduct, can be used to preserve market power, there is a separate question of how courts should make use of these theories. It is often difficult to distinguish harm to competition from harm to a rival, and this task is complicated by the fact that there can be an efficiency rationale for many types of exclusionary behavior.

Courts and Administration of Forced Sharing Rules

As stressed above, regulators have specialized expertise and fashion detailed rules that attempt to balance the competing interests of various groups of firms and consumers. In contrast, antitrust establishes general rules administered by nonspecialists. The Supreme Court’s decision in Trinko was based in part on its view that antitrust courts are poorly suited to establish and administer the terms of forced sharing requirements.

Forced sharing obligations generally require the establishment of prices and other terms of trade. As summarized above, the Supreme Court observed that courts typically lack sufficient industry knowledge or experience to do this properly. The problems faced by courts in administering the Modified Final Judgment that settled the government’s antitrust case against AT&T provide a stark reminder of problems that arise when courts

---

28 See, e.g., Ordover, Saloner, and Salop (1990), Whinston (1990), Bernheim and Whinston (1998), and Carlton and Waldman (2002). The application of the types of models discussed in these papers to antitrust is discussed in Carlton (2001).


take on an ongoing oversight function.” Judge Harold Greene approved the MFJ that settled the antitrust suit brought by the Department of Justice against AT&T and resulted in the breakup in 1984 of AT&T. Judge Greene’s oversight of the industry continued until the Telecommunications Act of 1996, which transferred Judge Greene’s remaining oversight responsibilities to the FCC. The MFJ, in effect, left one man responsible for implementation and ongoing supervision of the largest industrial restructuring undertaken in the history of the United States. Following the divestiture, Judge Greene had continuing responsibility for drawing, and redrawing, the lines that defined the types of services that the regional Bell operating companies (RBOCs) were permitted to offer and the conditions under which service could be offered. Between 1982 and 1996, Judge Greene issued thousands of advisory letters, received six thousand briefs, and often took years to resolve unopposed motions.

The adverse impact of inefficient administration of forced sharing rules is likely to be greatest in industries such as telecommunications, which are characterized by rapidly changing technology and innovation. New products and services are major sources of economic growth and increases in consumer welfare,” and a variety of recent studies shows that the introduction of new products such as cellular phones and later the Internet has resulted in enormous gains in consumer welfare.” Consumer harm from delays in the deployment of new services due to improper court oversight can defeat any desirable goals associated with impositions of duties to deal.

**Exceptions to Trinko**

It is important to emphasize that Trinko did not establish a blanket rule that eliminates in all circumstances the obligation under the antitrust laws to deal with competitors. The Court emphasized that “under certain circumstances, a refusal to cooperate with rivals can constitute anticompetitive conduct and violate Section 2.” As noted above, Verizon’s regulated status played an important role in the Court’s reasoning, and the Court stressed

“*This history is briefly retold in Thorne (2005), which provides a brief history of the administra-
tion of the MFJ, and the examples below are taken from this testimony. This history is told in more
detail in Huber et al. (2004, 2005).*

“Robert Solow’s (1957) classic study showed that technological innovations accounted for nearly seven-eighths of the growth in output per hour between 1909 and 1949. In his Nobel Lecture, Solow (1988, p. 313) concluded that his prior result “has held up surprisingly well in the thirty years since then during which time ‘growth accounting’ has been refined quite a lot...”

“Hausman (2004) estimated that cellular phones generated consumer surplus of between $80 billion and $150 billion in 2002 alone. On a per capita basis, this means that the average surplus per consumer is between roughly $600 and $1110 per year. Goolsbee and Klenow (2006) estimated that the Internet generates annual consumer surplus of $2500 to $3800 per consumer.

“540 U.S. 398, 408.
that "antitrust analysis must always be attuned to the particular structure and circumstances of the industry at issue.""

Still, the Court stressed that the circumstances that trigger the obligation to deal with one's rival under Section 2 are quite limited and referred to Aspen Ski, in which such obligations were imposed as "... at or near the outer boundary of Section 2 liability." The Court distinguished Trinko from Aspen Ski—Aspen Ski involved termination of an ongoing (and thus presumably profitable) joint venture, while, in Trinko, Verizon's regulatory obligation to open its network to CLECs was "something brand new." The Court emphasized the importance of this distinction.

From an economic perspective, this is a very thin reed on which to base a distinction. Changes in economic circumstances can convert an "ongoing" business relationship into "something brand new" requiring renegotiation between the parties. The imposition of a duty to deal with a rival in such circumstances may alter the bargaining positions of the parties, but the division of rents from a profitable venture is not a matter of antitrust concern. Court-imposed obligations to deal with rivals in such circumstances can impair economic efficiency by discouraging firms from entering into agreements in the first place and by interfering with ongoing economic relationships.

Although Trinko did not overrule Aspen Ski, the Court’s emphasis on the limited circumstances in which Aspen Ski applies makes it clear that an antitrust duty to deal will generally be difficult to establish. Similarly, the bipartisan Antitrust Modernization Commission recently concluded that "[r]efusals to deal... should rarely, if ever, be unlawful under antitrust law, even for a monopolist."

Antitrust Damages and Penalties for Regulatory Violations

Application of antitrust liability when firms fail to meet regulatory obligations to deal with rivals has the further effect of replacing the regulator-established system of penalties with treble damages under the antitrust

\[\text{Tbid. at 411.}\]
\[\text{Tbid. at 409.}\]
\[\text{Tbid. at 410. In discussing circumstances that trigger an obligation to deal with rivals, the Trinko opinion also mentions Otter Tail Power v. United States, 410 U.S. 366 (1973), and highlights that, in that matter, "the defendant was already in the business of providing a service to certain customers (power transmission over its network), and refused to provide the same service to other customers." Similarly, differences between new and existing customers can make a transaction with a new customer "something brand new."}\]
\[\text{See Carlton (2001) and Areeda (1990) for more extensive evaluations of the Court's decision in Aspen Ski.}\]
\[\text{Antitrust Modernization Commission (2007, p. 101). Carlton was a member of this commission.}\]
laws. As discussed above, the 1996 Act established a “carrot and stick” approach that penalized RBOCs that failed to meet their interconnection obligations by preventing them from obtaining permission to provide long-distance service. In addition, state regulations established detailed interconnection requirements and penalties for failure to comply with these requirements.

While both regulation and antitrust laws embody penalties for violations, the circumstances at issue in *Trinko* leave little economic rationale for the imposition of potentially enormous antitrust penalties on top of the specified penalties established by regulators. More specifically, Verizon was subject to a variety of regulatory obligations, including Verizon’s interconnection agreement with AT&T (the CLEC from which Trinko wished to obtain service), industrywide guidelines agreed to by Verizon, the CLECs, and the New York Public Service Commission, as well as the “Performance Assurance Plan” that established detailed measures of Verizon’s performance in providing service to CLECs and established penalties for Verizon’s failure to meet specified standards. Under the plan, the size of the penalty depended on the magnitude of the performance shortfall, Verizon’s historical performance in meeting this standard, and the impact on the CLEC’s business.43

The transformation of Verizon’s failure to provide sufficient assistance to rivals into an antitrust violation would have added treble damages suffered by plaintiffs as an additional penalty. Where there is no antitrust liability in the absence of regulation, it follows that the public interest goal can be adversely affected by making regulatory violations triggers for additional antitrust penalties. For example, the risk of very large antitrust penalties can lead to “excessive” cooperation by incumbent firms with rivals, which in turn can harm consumer welfare by discouraging investment, innovation, and competition between ILECs and CLECs.

Moreover, the traditional economic rationale for trebling damages under the antitrust laws does not apply in *Trinko*, or likely in a variety of related settings. The trebling of antitrust damages is typically justified by the difficulty of detecting violations—if violations are difficult to detect, then establishment of the appropriate incentives for deterring violations requires that violators once detected face penalties that exceed the harm to the plaintiff, and trebling is one way to accomplish this.44 However, the extensive monitoring of Verizon’s performance by state regulators and by CLECs and the ability to compare Verizon’s performance with that of other ILECs implies that violations by Verizon or other ILECs will not go undetected. There is little economic rationale for adding trebled antitrust penalties to those of regulators under these circumstances.

43 Some aspects of the regulatory oversight and penalties faced by Verizon are summarized in 540 U.S. 398 412-3 (2004). This regulatory framework is described in more detail in Thome (2005).
ADDITIONAL ISSUES RAISED IN
THE TRINKO OPINION

Verizon v. Trinko and U.S. v. ÄRLT

Several observers have noted the similarities in the economic issues raised in Trinko and those in the Justice Department’s antitrust suit against AT&T (which was settled with the MFJ in 1982), as well as the difference in the outcomes.  The economic foundation of both cases is the well-recognized principle that a local exchange monopolist facing a regulated retail price (presumably set below the profit-maximizing level) has an incentive to discriminate against rival firms that can purchase network access from the monopolist at a regulated input price (set below the profit-maximizing level) that is so low that it allows the rival to price profitably at or below the regulated retail price. In the Department of Justice case, it was alleged that AT&T—then the incumbent ILEC—denied network access to rivals in order to exclude competition for its long-distance business; in Trinko, it was alleged that Bell Atlantic denied access to rivals in order to exclude competition for its local telephone business.

While U.S. v. AT&T settled before going to a complete trial and appellate review, the discrimination-based antitrust claims resulted in a massive industrial restructuring and resulted in the creation of separate firms providing local telephone service and long-distance service. In Trinko, of course, the Court failed even to recognize that discrimination against rivals by an incumbent monopolist was sufficient to establish an antitrust claim. The outcomes of the two cases reflect divergent views about the comparative advantages of antitrust and regulation in addressing competitive problems.

Both the Department of Justice’s allegations in U.S. v. AT&T and the MFJ itself reflect a distrust of (1) the ability of regulators to design and implement rules that foster entry into the segments of the telecommunications industry that could sustain competition and (2) the ability of regulators to identify and penalize discrimination by an incumbent monopolist against rivals attempting to obtain network access.” If regulators are unable to perform these functions adequately, then discrimination incentives can be eliminated only through structural remedies and prohibitions on ILEC activities such as those embodied in the MFJ.

Time has demonstrated that there are significant costs associated with the MFJ’s divestiture remedy. Divestiture and the line-of-business restrictions in the MFJ precluded the RBOCs from offering a variety of services for which they were likely efficient providers and eliminated their ability to

---

realize likely economies of scope. While the divestiture and the MFJ circumvented a number of concerns about the efficacy of regulation, they introduced a number of new ones and put in place an extraordinarily cumbersome mechanism for administering the settlement. The 1996 Telecommunications Act introduced new problems. Regulatory issues relating to the pricing of network elements have been disputed almost continuously both before and since passage of the 1996 Act.

There is little doubt that the Court’s decision in *Trinko* was shaped by this well-known history, with the *Trinko* decision highlighting the limitations faced by generalist judges and courts in these circumstances. More generally, the Court in *Trinko* also had a more favorable view of the ability of regulators to detect and remedy discrimination compared with the views held by the Department of Justice at the time of *U.S. v. AT&T*. For example, the Court’s opinion in *Trinko* describes at considerable length how "[t]he regulatory framework that exists in this case demonstrates how, in certain circumstances, 'regulation significantly diminishes the likelihood of major antitrust harm.'" More specifically, the Court favorably cites the success of the FCC’s process for authorizing RBOC entry into the provision of long-distance service, the New York Public Utility Commission’s Performance Assurance Plan, and concludes that "[t]he regulatory response to the OSS failure complained of in respondent’s suit provides a vivid example of how the regulatory regime operates."

Reconciling *Trinko* with the 1996 Act’s "Savings" Clause

As noted above, the 1996 Act has an antitrust "savings" clause that states that "nothing in this Act [...] shall be construed to modify, impair, or supersede the applicability of any of the antitrust laws." The complaint in *Trinko* relies on this clause to argue that regulatory violations that raise price and reduce output relative to the level that would have existed in the absence of the violation are anticompetitive and thus are appropriately interpreted as antitrust violations.

The Supreme Court quickly dismissed this approach and concluded that the savings clause in the 1996 Act "does not create new claims that go beyond existing antitrust standards" and evaluated Trinko’s claims based on existing antitrust standards regarding a firm’s obligation to deal with its rivals. Trinko’s argument, if accepted by the Court, would have transformed many regulatory violations into antitrust offenses.

Does the Court’s opinion here leave any role for antitrust in regulated industries? Yes, albeit one that is more limited than the simple application of an economic test of assessing antitrust liability based on considerations

47 540 U.S. 398, 412.
48 Ibid. at 413.
of price and output alone. Even in regulated industries, many activities remain unregulated, and antitrust violations can arise in a variety of activities outside the scope of a regulated firm's failure to comply with regulatory obligations. For example, although the FAA's regulatory authority regarding airline safety gives it the ability to influence how aircraft are maintained, that would not immunize the airlines from antitrust liability if they engaged in price fixing. In addition, regulators often rely on markets instead of specific rules to govern some aspects of regulated firms' activities, and the Court's decision continues to subject these activities to the antitrust scrutiny that exists in unregulated sectors.

The recommendation adopted by the Antitrust Modernization Commission fits closely to the rules applied by the Supreme Court for defining the scope of antitrust in regulated industries. The commission recommended:

> When the government decides to adopt economic regulation, antitrust law should continue to apply to the maximum extent possible, consistent with that regulatory scheme. In particular, antitrust should apply wherever regulation relies on the presence of competition or the operation of market forces to achieve competitive goals.

**CONCLUSION**

The Supreme Court's unanimous decision in *Trinko* relies heavily on economic reasoning in drawing a sharp line distinguishing a firm's regulatory obligation to deal with rivals from any such obligation that might arise under the antitrust laws. The Court reasoned that the imposition of a general obligation to deal with rivals can harm social welfare by discouraging a firm's incentives to invest, innovate, and, more generally, compete. The Court also concluded that courts are typically at a significant informational disadvantage relative to regulators in addressing the pricing and other terms of trade that necessarily arise from imposition on a monopolist of a duty to deal with rival firms. The Court explained that only in rare circumstances does there arise a duty to deal with rivals under the antitrust laws. The Court refused to expand those circumstances when a firm is regulated.

**REFERENCES**


CASE 19


Robert S. Pindyck

INTRODUCTION

Payment card networks provide what has become essential infrastructure for much of the country’s economic activity, so it is important that they compete effectively and operate efficiently. If two networks restrict the ability of their members to join other networks, should that raise antitrust concerns? And what if those two networks also have shared governance? Could such practices have pro-competitive aspects that benefit consumers, or should they always be prohibited?

In October 1998, the Antitrust Division of the U.S. Department of Justice brought an antitrust suit against two of the major payment card networks, Visa and MasterCard. Any bank that issues Visa cards is free also to issue MasterCard cards, and vice versa, and indeed, many banks issue the cards of both networks. However, both Visa and MasterCard had rules prohibiting their issuing banks from also issuing the cards of other networks, in particular American Express and Discover. In addition, most banks are members of both the Visa and MasterCard networks, and it is the members that control the networks, so that governance of the two networks is to

*My thanks to the editors, John Kwoka and Larry White, as well as Dennis Carlton, Emily Cotton, David Evans, Steve Herscovici, Rebecca Kirk, Martha Samuelson, Richard Schmalensee, and Rebecca Weiss for helpful comments and suggestions. The author served as an expert witness for MasterCard in its antitrust case against the Government; the views expressed here are solely those of the author and should not be attributed in any way to MasterCard.


Page 507
some extent shared. The DOJ claimed that both practices—restrictions on issuance and shared governance—were anticompetitive and should be prohibited. Visa and MasterCard argued that these practices were, on net, pro-competitive. In the end, after a trial and a failed appeal by the two card networks, the DOJ won on its issuance restriction claim but lost on shared governance.

Putting the results of the trial aside, the Government’s case made several things clear: The structure of the payment card industry is unique, market definition is unclear, the nature of competition among card networks is complex, and assessing the competitive impacts of various practices is not easy. The case also raised fundamental questions. Most importantly, given that many banks issue both Visa and MasterCard credit cards and that most merchants that accept one also accept the other, do the two networks really compete, and if so, how? And do Visa and/or MasterCard have market power and, if so, in what market, and how is it exercised?

To understand the issues in this case, one must first understand how payment card networks operate, differ from each other, and compete. To this end, the next section provides a brief overview of the structure of the industry. I return to the DOJ’s case against Visa and MasterCard in the following section and discuss the Government’s claims in more detail. The next section deals with the question of market definition and contrasts the DOJ’s view to that of the defendants. The sections that follow discuss the competitive effects of the Visa and MasterCard issuance restrictions and the competitive effects of shared governance, once again contrasting the DOJ’s and the defendants’ views. The remaining section discusses the outcome of the case and offers some concluding remarks.

INDUSTRY STRUCTURE

Before turning to the Government’s antitrust case and trying to assess the competitive impacts of the specific practices at issue, it is important to be clear about how payment card networks operate, the nature of the product or service they provide, and how consumers and merchants view them in relation to other products and services. Accordingly, a brief overview of the structure of the industry is appropriate.

The Government’s case was wholly separate from a case against the two networks that was brought by Wal-Mart and other large merchants, alleging illegal tying of debit cards to credit cards: In re Visa Check/MasterMoney Antitrust Litigation (Case No. 96-CV-5238), U.S. District Court for the Eastern District of New York. The merchants claimed that the associations’ "Honor All Cards" rules, which required merchants that accepted Visa and MasterCard credit cards also to accept Visa and MasterCard signature debit cards, allowed the associations to charge supracompetitive interchange rates on debit transactions. The case was settled in 2003; the associations agreed to pay the plaintiff class more than $3 billion and to separate acceptance of debit cards from acceptance of credit cards.

For a thorough and very readable account of the structure and history of the payment card industry, see Evans and Schmalensee (2005). For a discussion of some of the underlying economic theory, see Rochet and Tirole (2002).
What Are Payment Cards, and What Do They Do?

General-purpose payment cards come in three basic flavors: charge cards, credit cards, and debit cards. What all three types of cards have in common is that they act as a payment mechanism; that is, they provide an essential service: a means of payment. As long as a merchant accepts a particular payment card, one can use it to make purchases from that merchant. Of course, payment cards are not the only products that act as a payment mechanism; cash, checks, Internet payment accounts (such as PayPal), and stored value cards, among other products, do as well. Indeed, an important question, related to market definition and discussed below, is the extent to which these various payment mechanisms are substitutable with each other.

How do these three types of payment cards differ? A charge card provides a means of payment and, other than some ancillary services (such as rental car and travel insurance, end-of-year expenditure summaries, etc.), only a means of payment. In terms of annual charge volume, the leading charge card in the United States is the American Express (Amex) card, which in 2006 had an annual U.S. charge volume of about $400 billion. (American Express also issues credit cards, but that has historically been a relatively small part of its business.) One can use an American Express charge card to pay for purchases from merchants that accept the card but not to borrow money; that is, the cardholder is expected to pay each monthly balance in full.

A credit card provides a means of payment, bundled with a second service, revolving credit. Of consumers who hold credit cards, about half regularly make use of the credit service (i.e., they maintain revolving balances), and the remainder do so rarely or never. Finally, a debit card also provides a means of payment, but it is linked to the cardholder’s checking account, from which purchase costs are debited as they are incurred. At the discretion of the cardholder’s bank, the checking account might include a line of credit, giving the debit card a credit-like feature. Some debit cards are essentially bank ATM cards that can be used to make purchases by entering a PIN number after swiping the card through a card reader (“online” debit). Most debit cards, however, are issued within the Visa or MasterCard networks, have the Visa or MasterCard logos, and can be used to make purchases in the same way as a credit card (“offline” or “signature-based” debit) or, at the discretion of the cardholder, “online” by entering a PIN number.

Of the three types of payment cards, the most rapid growth in the United States recently has been in the use of debit cards. Table 19-1 shows debit purchase volume and total credit plus debit purchase volume for the years 2000 through 2006. Over this six-year period, debit purchase volume has grown at an annual rate of about 21 percent, from about 16 percent of total purchase volume in 2000 to about 27 percent in 2006.

The Government’s antitrust case focused almost entirely on the use of cards as a means of payment. Thus, the discussion that follows will likewise
focus on the payment mechanism service, which is provided by all three types of cards.

What Is a Card Network?

A card network consists of a set of consumers who hold and use the card, a set of merchants that accept the card as a means of payment, a system for processing payment flows, and a set of rules or standards governing the design, uses, and acceptance of the card. Any card issued through the MasterCard network, for example, must show the MasterCard logo, and any merchant that accepts MasterCard must also display the logo and may not impose a surcharge on a consumer for using the card instead of some other form of payment. Thus, the cardholder knows that her card can be used interchangeably with cash at no additional cost. (This last point is particularly important for the issue of market definition.)

Clearly the value to consumers and merchants of participating in a payment card network increases with the size of the network, that is, with the total number of consumers who carry and use the cards, and with the number of merchants that accept the cards. This creates a “chicken and egg” problem: To create a card network, consumers must be convinced that merchants will accept the card, and merchants must be convinced that

| TABLE 19-1
Growth of Debit and General Purpose Card Use |
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Debit Cards</td>
</tr>
<tr>
<td>Purchase Volume (billion)</td>
</tr>
<tr>
<td>[A] 2000</td>
</tr>
<tr>
<td>[B] 2001</td>
</tr>
<tr>
<td>[C] 2002</td>
</tr>
<tr>
<td>[D] 2003</td>
</tr>
<tr>
<td>[E] 2004</td>
</tr>
<tr>
<td>[F] 2005</td>
</tr>
<tr>
<td>[G] 2006</td>
</tr>
</tbody>
</table>

Sources:
consumers hold the card and want to use it. This makes the creation of a new network difficult at best and creates challenges for an existing network that wants to expand to other regions of the world where currently its presence is minimal or nonexistent.

Every network must deal with three crucial functions—issuing cards to consumers, signing up (“acquiring”) merchants and servicing them, and processing transactions. These functions are performed in different ways, depending on whether the network is "open" or "closed."

Open versus Closed Networks
Visa and MasterCard are open networks. As long as it is willing to abide by the network’s rules and standards, any bank can join the Visa or MasterCard associations and issue cards to consumers under the Visa or MasterCard networks. Likewise, any bank can become a Visa or MasterCard “acquirer” and sign up and service merchants. Issuing banks benefit from annual fees charged to cardholders, from the fees and interest payments from cardholders who maintain revolving balances, and from the “interchange” fee collected on all card transactions. Acquiring banks likewise receive (much smaller) fees on transactions. The Visa and MasterCard associations provide authorization and settlement services (often through contracts with third-party processors), for which they charge small fees, and also set the networks’ rules and standards, advertise the network brands, and support product innovation (such as the development of "smart cards"). Issuing banks will also innovate and design products (such as "affinity" cards that are linked with some organization, e.g., a university alumni credit card) and are free to set their fees and the interest rates that they charge on credit balances. Visa and MasterCard were operated as not-for-profit associations of their respective issuing and acquiring member banks, and they have been governed by and operate for the benefit of their member banks.

By contrast, through 2004, American Express and Discover were closed networks. Only American Express issued American Express charge cards to consumers in the United States and acquired and serviced merchants. American Express would also process all transactions and invest in

---

4 Payment card networks are an example of a "two-sided platform," that is, an entity that brings together two or more different groups of customers, in this case cardholders and merchants. For a general discussion of two-sided platforms, including antitrust issues that sometimes arise, see Evans (2003) and Evans and Schmalensee (2008).

5 Visa and MasterCard were associations of member banks at the time that the Government filed its antitrust claims. Since then, MasterCard has become a publicly traded corporation, and in October 2006 Visa announced its intention also to become a publicly traded corporation. However, MasterCard and Visa remain open networks and operate largely in the interests of the member banks (which in the case of MasterCard hold large equity shares).

6 This is no longer true. American Express decided in 1996 to open its network and seek bank issuers, but given the cost of losing its Visa and/or MasterCard portfolio, no bank concluded an
advertising and innovation to promote and expand the American Express network. Discover has been doing the same for its network.

In a Visa or MasterCard credit card purchase, the merchant receives about 98 percent of the amount of the purchase. The remaining 2 percent is called the "merchant discount," which is the sum of the fee paid to the merchant's acquiring bank, the Visa or MasterCard processing fee, and the roughly 1.5 to 1.7 percent interchange fee that is collected by the cardholder's issuing bank. Thus, the interchange fee, which is set by the Visa or MasterCard organization, accounts for most of the merchant discount. It is important to note that the issuing bank owns the consumer's account and takes the payment risk. The interchange fee helps to offset that risk as well as other costs of issuance. In an Amex purchase, on the other hand, American Express—which owns the consumer's account and thus takes the payment risk—receives the entire merchant discount (which for most transactions has historically been about 2.5 to 3 percent).

Open and closed networks each have their own advantages and disadvantages. As a closed network, American Express has more control over its brand and can better maintain uniform standards when dealing with cardholder requests and complaints and responding to fraudulent transactions. In addition, it can collect and utilize data on all its cardholders, making it more effective in "cross-selling" other products and services, such as travel agency and concierge services. In an open network, cardholders deal with their issuers, and issuers will differ in terms of their responsiveness to cardholder complaints and overall quality of service. On the other hand, issuers in an open network compete with each other to attract consumers, which can lead to more rapid network expansion. Also, various issuers may find different ways to innovate in terms of the products and services they offer.

**Duality**

Most banks that have been members of the Visa association have also been members of the MasterCard association, and many banks issue the cards of both networks. Because each association is essentially controlled by its members, overlapping membership also implies overlapping governance. Members of the two associations can sit on the boards of either one (but not issuance deal with Amex. However, as an outcome of the Government's antitrust suit, Visa and MasterCard can no longer prevent their member banks from issuing American Express or Discover cards. At the time of this writing, American Express had negotiated issuance agreements with six banks. More on this later in the discussion of the outcome of the case.

Visa, MasterCard, American Express, and Discover are the major payment card networks in the United States. The focus will be on them, and the much smaller networks, such as Diners Club and JCB, will be largely ignored.
on the boards of both). This overlapping membership is generally referred to as "duality." However, it is important to distinguish between two aspects of duality: issuance duality refers to the fact that many banks issue the cards of both networks, while governance duality refers to the shared governance of the two associations. The DOJ objected to governance duality, not issuance duality (although as a practical matter it is unclear how the two could be separated).

In addition to issuance duality, almost every U.S. merchant that accepts Visa also accepts MasterCard, and vice versa. This overlapping merchant acceptance, overlapping membership, and overlapping governance raise two questions.

First, do we really need both networks? Isn't MasterCard (or Visa, the larger network) redundant, or would consumers and merchants lose if one of the networks were to disappear? Second, even if we accept that the two networks compete with each other, thereby creating benefits from having both of them, isn't governance duality anticompetitive in that it must reduce the intensity of competition?

Neither the Government nor Visa or MasterCard claimed that there is redundancy; all parties agreed that there is substantial competition between the networks, so there is a benefit from having both of them in operation. Banks that are members of both networks benefit from this competition because even though it consumes resources (which are ultimately a cost to the member banks), it leads to innovation and network expansion, which are beneficial to the members. The Government claimed, however, that governance duality harms consumers by reducing the intensity of the competition between the two networks and should therefore be prohibited. To assess this claim, one must understand how payment card networks compete.

**Competition among Networks**

Competition among Visa, MasterCard, Amex, and Discover occurs through the merchant discount, advertising, and innovation. The main objectives of competition are the expansion of the network itself and the wider usage of the network's cards. All four organizations promote the expansion of their networks into new retail and service sectors through incentive interchange and discount rates and promotional programs. For example, in the early 1990s all of the networks encouraged supermarkets and grocery stores to accept their cards by offering those stores a lower merchant discount. (Visa and MasterCard did this by reducing the interchange fee, and Amex by directly reducing the merchant discount, on purchases from supermarkets and grocery stores.) More recently, Visa and MasterCard successfully expanded into smaller transactions, such as "swipe-and-go" (no signature required) purchases at McDonald's.
The networks also compete for new and existing cardholders via cardholder fees and rewards programs, which also encourage cardholders to use their cards at every opportunity, and via brand advertising, technological and product innovations, and promotional deals to acquire new accounts. Advertising is important because it creates brand identity, which helps expand the network. There is also some network differentiation. American Express has targeted higher-income consumers whose average purchases are larger, and as a result, merchants are willing to accept a higher merchant discount. Discover has traditionally targeted the other end of the income spectrum. Visa and MasterCard have been in the middle, although more recently they have sought higher-income consumers through the development and promotion of their premium cards.

Competition between the Visa and MasterCard networks also occurs via advertising, innovation, and the development and introduction of new types of cards (such as affinity and co-branded cards). But, in addition, these two networks compete through, and for, the issuing banks. The idea of competing through issuing banks may seem counterintuitive, given duality (i.e., the fact that most issuing banks belong to both networks), but duality can actually enhance competition. The reason is that by belonging to both networks, member banks can switch part or all of their cardholder portfolios from one network to the other and use this ability to extract concessions from the networks in the form of cash payments, assistance with technology adoption, and so on.

Competition also occurs among issuers, which compete to sign up (and retain) cardholders. This competition takes the form of consumer solicitations, offering more advantageous terms (e.g., with respect to interest rates, annual fees, rewards, and rebates), and introducing new products such as affinity cards (where rebates might go to, say, the cardholder’s college alma mater). Competition has increased over the years as issuers try to differentiate their cards from each other by offering such benefits as frequent flier miles, rebates on purchases, no or low annual fees, purchase protection/security, discounted introductory interest rates, and travel accident insurance. Issuer competition has helped to expand the Visa and MasterCard networks and thus is a benefit of an open network.

In 2006, total U.S. payment card charge volume for purchases and cash advances for the four major networks, including signature-based debit cards, was about $3.0 trillion, up from about $2.1 trillion in 2003. Table 19-2 shows the 2006 market shares of the four major networks. Of the 2006 total, Visa accounted for 53.6 percent (about $877 billion in credit and $720 billion in debit), MasterCard accounted for 28.9 percent (about $610 million in credit and $252 million in debit), American Express accounted for 13.7 percent, and Discover 3.8 percent. Excluding debit, the Amex share was about 20 percent, and the Discover share about 6 percent.
Although the Government agreed that Visa and MasterCard compete with each other and with American Express and Discover, it claimed that governance duality and the issuance restrictions reduced the intensity of that competition and that, as a result, Visa and MasterCard were in violation, on two counts, of Section 1 of the Sherman Antitrust Act. Specifically, the DOJ stated in its Complaint that it was acting to prevent Visa and MasterCard from "violating the antitrust laws by restraining competition in general purpose card network products and services." The Complaint alleged that the governance structures of the two associations "... substantially lessened competition between Visa and MasterCard" and that the membership rules and policies of the associations have "[e]liminated] certain forms of competition among the Visa and MasterCard member banks, and have effectively precluded American Express and Discover/Novus from competing to enlist banks in the United States to issue their cards." The Complaint also alleged that these Section 1 violations occurred in the context of a relevant market consisting of "[t]he products and services provided by general purpose card networks," which, interestingly, excluded debit cards and other forms of payment.


*Complaint, 114. The complaint claims (18) that "General purpose cards are payment devices that a consumer can use to make purchases (a) from unrelated merchants and (b) without accessing or

Table 19-2
Market Shares by Total Charge Volume ($billion)

<table>
<thead>
<tr>
<th></th>
<th>2006</th>
<th>Percentage of Total</th>
<th>Excluding Debit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visa Credit</td>
<td>1596.44</td>
<td>53.6%</td>
<td>43.7%</td>
</tr>
<tr>
<td>Visa Debit</td>
<td>876.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MasterCard</td>
<td>719.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MasterCard Credit</td>
<td>862.35</td>
<td>28.9%</td>
<td>30.4%</td>
</tr>
<tr>
<td>MasterCard Debit</td>
<td>610.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>American Express Credit</td>
<td>406.80</td>
<td>13.7%</td>
<td>20.3%</td>
</tr>
<tr>
<td>American Express Debit</td>
<td>252.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discover Credit</td>
<td>113.55</td>
<td>3.8%</td>
<td>5.7%</td>
</tr>
<tr>
<td>Total Signature-Based Cards</td>
<td>2979.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Cards Excluding Debit</td>
<td>2007.33</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As the plaintiff, to prove these claims and show competitive harm the DOJ had to demonstrate five things: (1) its definition of a "general purpose card" market—as opposed to some broader market—was consistent with the data; (2) Visa and MasterCard should be treated as a single entity that exercised monopoly power or that Visa and/or MasterCard individually had and exercised monopoly power; (3) governance duality reduced competition between Visa and MasterCard and thereby contributed to their monopoly power; (4) Visa and MasterCard exercised monopoly power through their issuance restrictions, and this resulted in harm to Amex and/or Discover; and (5) the exercising of monopoly power by Visa and MasterCard resulted in a market outcome significantly different from a competitive outcome, harming consumers.

The objective of the antitrust laws is to protect consumers and not necessarily competitors. Suppose that dual governance and issuance restrictions did reduce competition between Visa and MasterCard and did enhance their market power by weakening American Express and Discover. How would that harm consumers? The DOJ alleged that dual governance reduced the incentives for Visa and MasterCard to compete by constraining innovation and investments in new and better products, resulting in the delayed introduction of chip-based smart cards, Internet encryption standards, informative advertising, and premium cards. The DOJ also alleged that the Visa and MasterCard issuance restrictions limited consumer acceptance of Amex and Discover cards, thereby reducing merchant acceptance levels and restraining the ability of American Express and Discover to develop and distribute new products and features such as smart cards. The DOJ further alleged that the issuance restrictions foreclosed American Express or Discover from issuing offline debit cards (which in the future could be linked to credit card functions on a single relationship card with encrypted technology) and deprived consumers of the ability to obtain credit cards that combine the features of their preferred bank with any of four network brands, each of which has different qualities, characteristics, features, and reputations.

The case went to trial on June 12, 2000, before U.S. District Judge Barbara S. Jones.10 Judge Jones’s decision was filed on October 9, 2001, and after an appeal by Visa and MasterCard, was affirmed by the Second Circuit Court of Appeals on September 17, 2003.11 The associations’ appeals were exhausted when the Supreme Court denied certiori in October 2004.

10 Reserving the consumer’s funds at the time of purchase,” which would include credit and charge cards but not debit cards.

11 Expert testimony was presented by Michael Katz on behalf of the DOJ, by Richard Schmalensee and Richard Rapp on behalf of Visa, and by Robert Pindyck on behalf of MasterCard.

MARKET DEFINITION

To determine whether Visa or MasterCard had market power, one must begin with a definition of the relevant market. Clearly the credit and charge cards issued by Visa, MasterCard, American Express, and Discover are in the market, but are there other products that act as payment mechanisms and that are reasonably close substitutes with credit and charge cards? Should debit cards, for example, or other means of payment, such as cash and checks be included in the market? And should supply substitutability, as well as demand substitutability, be part of market definition?

The textbook definition of a market is the collection of buyers and sellers that, through their interactions, determine the price of a product or set of products. Clearly, then, both supply and demand substitutability determine market boundaries. Two products will be in the same market if either demand substitutability or supply substitutability is high because either condition will lead to a high correlation of prices for the two products. U.S. antitrust authorities, however, typically focus only on demand substitutability and limit the supply-side analysis to the possibility of entry (see, e.g., the discussion of market definition in the DOJ’s Merger Guidelines), and indeed Judge Jones did just that in her decision. But in fact both demand and supply-side substitutability are relevant and are discussed below.

One difficulty faced by both sides of this case is that there is no clear “price” that users of payments cards (versus other means of payment) pay, and this complicates any assessment of demand or supply substitutability based on the estimated impacts of a price change. For payment cards, the price that consumers pay might be some combination of annual fees, late fees, and other financing charges, minus the value of rewards and rebates. The price that merchants pay—which is effectively the price received by the parties that make up the card network—is the merchant discount. What about the price of an alternative means of payment such as cash? For consumers the price is the inconvenience of obtaining and carrying cash, including the possibility of loss or theft. For merchants, the price is largely the cost of theft, which can be substantial but may be hard to quantify.

All parties agreed that the relevant geographic market was the United States. The disagreement was over the set of products that should be included in the market. As one might expect, the DOJ’s position was quite different from that of Visa and MasterCard.

*See, for example, Pindyck and Rubinfeld (2005).*

*Suppose, for example, that one is trying to decide whether two airline routes—New York to San Francisco and New York to Los Angeles—are in the same market. Demand substitutability across these routes is probably low. However, insofar as airlines serving those routes could readily shift flights from one route to the other in response to price and profit differences (so that supply substitutability is high), both routes will be in the same market.*
The Government's Position on Market Definition

As mentioned above, the DOJ claimed that relevant market consisted of "[t]he products and services provided by general purpose card networks."

The DOJ argued that credit and charge cards issued through the Visa, MasterCard, American Express, and Discover networks are included in the market under this definition, as well as the Diners Club charge card (which is used by relatively few consumers). According to the DOJ, however, debit cards—whether cards in the Visa or MasterCard networks that can be used offline or online cards—should be excluded from the market. Other forms of payment, such as cash and checks, were likewise excluded by the DOJ.

The DOJ based its definition of the relevant market on a view about demand substitutability that focused on the notion that general-purpose cards provide specific functions that other products do not provide. The specific functions cited by the DOJ include widespread acceptance, security, and deferred payment options. The DOJ argued that other forms of payment do not have all of these functions and thus belong in a separate market. (Note, however, that charge cards generally do not provide deferred payment options beyond the grace period, whereas checks with overdraft protection do.) The exclusion of debit cards was based on the notion that they do not provide a deferred payment option (although they would if the bank provides overdraft protection). The DOJ also ignored the considerable supply substitutability that exists between credit and debit; credit and debit use the same network facilities, so that many banks can easily choose to issue either one or both.

The DOJ’s market definition implied that Visa and MasterCard had a large combined market share and even large individual shares. Based on the DOJ’s market definition, at the time of the trial Visa’s share of payment card purchase volume in the United States was about 45 percent and MasterCard’s was about 30 percent, for a combined market share of about 75 percent. Both the individual and combined shares were sufficiently large for the DOJ to argue that Visa and MasterCard had monopoly power.

Visa and MasterCard Positions on Market Definition

Visa and MasterCard argued that the relevant market was much broader than just credit and charge cards—that it certainly should include debit cards, as well as other general forms of payment such as cash and checks. They pointed out that it can be misleading to define markets in terms of functionality because functionality is a vague concept, it often has many dimensions, and it is hard to measure. Instead, markets are usually defined based on demand and/or supply substitutability. Focusing on demand substitutability, one must ask whether credit and charge cards are substitutable with other forms of payment for at least a reasonable fraction of purchasing decisions by consumers.
Visa and MasterCard argued that from the point of view of consumers, although they are not perfect substitutes, credit and charge cards, debit cards, cash, and checks are highly interchangeable, at least for many purchases. Although credit cards and cash are probably not interchangeable when it comes to paying a large hotel bill, they are quite interchangeable for smaller purchases, and indeed many consumers will sometimes pay with a credit card and sometimes with cash (or a check). MasterCard supported this argument in part by conducting a telephone survey of consumers, the results of which showed a considerable willingness to substitute among different forms of payment. As for debit cards, the offline cards that are part of the Visa or MasterCard networks function much like credit and charge cards and are virtually indistinguishable from Visa and MasterCard credit cards in terms of their use and acceptance. In addition, merchants often view these different means of payment as highly interchangeable, as they all have fairly similar costs associated with them. (Although there is no merchant discount for cash or checks, as there is for credit, charge, and debit cards, there are other costs of accepting cash or checks, such as the risk of theft or a check bouncing.) Finally, as mentioned above, Visa and MasterCard argued that the considerable supply substitutability between credit cards and debit cards also implies that they are part of the same market.

Broadening the market definition to include offline debit does not change the Visa and MasterCard shares much because most offline debit is within the Visa and MasterCard networks. However, if cash and checks are also included in the market, the combined Visa-MasterCard share falls dramatically.

The Court’s Opinion on Market Definition and Market Power

Judge Jones adopted the DOJ’s market definition: There is a general-purpose card market separate from other forms of payment. It seems that this decision was arrived at in part because of the difficulty of measuring demand substitutability. The decision stated that “it is highly unlikely that there would be enough cardholder switching away from credit and charge cards to make any such price increase unprofitable for a hypothetical monopolist of general purpose card products. This conclusion is buttressed by the fact that (1) few, if any, cardholders actually can or do observe price increases, including interchange rate increases and increases in service fees charged by issuing banks; and (2) the burden of such increases is at least partly passed on by merchants and so is shared by consumers who use other means of payment.”

With respect to debit cards, the court decided that "Consumers also do not consider debit cards to be substitutes for general purpose cards." This view was not surprising for online debit cards, which, "due to their relative lack of merchant acceptance, their largely regional scope, and . . . which require a pin number, are not adequate substitutes for general purpose cards." For offline debit, on the other hand, the court cited Visa and MasterCard research, which "demonstrates that consumers do not consider off-line debit cards to be an adequate substitute for general purpose cards, even though they have attained widespread merchant acceptance."

With this rather narrow definition of the market, the court came to the conclusion that Visa and MasterCard each individually had market power. The evidence that the court cited to support this conclusion was examples of conduct. For example, the decision mentioned that "both Visa and MasterCard have recently raised interchange rates charged to merchants a number of times, without losing a single merchant customer as a result." And: "Defendants' ability to price discriminate also illustrates their market power. Both Visa and MasterCard charge differing interchange fees based, in part, on the degree to which a given merchant category needs to accept general purpose cards."

By coming to the conclusion that the defendants had market power, the court could proceed to evaluating the specific practices that the DOJ claimed were Sherman Act Section 2 violations—issuance restrictions and dual governance.

COMPETITIVE EFFECTS OF ISSUANCE RESTRICTIONS

Visa's Bylaw 2.10(e) precluded its member banks from issuing the credit and charge cards of any competing network other than MasterCard's. Likewise, MasterCard's Competitive Programs Policy precluded its member banks from issuing the credit and charge cards of any competing network other than Visa's. Both Visa and MasterCard instituted these regulations to help maintain the loyalty of their members while preventing individual members from free-riding on the assets of the respective association.

The DOJ claimed that these rules prevented American Express and Discover from gaining access to banks as potential issuers and thus prevented them from competing as effectively as they could have. In particular, the issuance restrictions reduced the number of consumers who carry

*United States v. Visa USA Inc., at 336.
"MasterCard's membership policy grandfathered two arrangements: Citibank's ownership of the Diners Club network and Household Bank's plan to issue JCB cards.
Amex and Discover cards, thereby reducing merchant acceptance levels and restraining the ability of American Express and Discover to innovate and develop new features, with the overall result of decreased network-level competition and fewer and less-varied credit card products available to consumers. The DOJ argued that consumers were also harmed because they were unable to obtain credit cards that combine the features of their preferred bank with any of four network brands, each of which has different characteristics and features.

**Anticompetitive Effects of Issuance Restrictions**

Visa’s Bylaw 2.10(e) and MasterCard’s Competitive Programs Policy did not foreclose American Express and Discover from the payment card industry: indeed, for American Express merchant acceptance, the number of cardholders, and annual charge volume had all grown substantially during the decade preceding the antitrust suit. On the other hand, these policies did prevent American Express and Discover from issuing their cards through most banks in the United States unless those banks were willing to cease issuing the cards of Visa and MasterCard.

Recall that one of the advantages of an open network like Visa’s is that issuers compete with each other to attract consumers. In addition to advertising and promotional mailings, issuers compete by finding different ways to innovate and differentiate the products and services they offer. This innovation and differentiation can take the form of designing and offering new products (such as premium cards and co-branded or affinity programs with airlines and other organizations) and by setting different fees and interest rates. By promoting their products and enabling consumers to find products that more closely match their preferences, this issuer competition may help to expand the size of the network, which in turn makes it more valuable to consumers and merchants alike. Of course, American Express and Discover also can (and do) promote their networks through advertising and mailings and innovate and develop new products and services. But the Visa and MasterCard issuance restrictions forced them to do this largely on their own. To the extent that this reduced Amex’s and Discover’s investments in marketing, innovation, and product design, it led to less-diverse product offerings and networks that were smaller than they would have been otherwise, thereby reducing consumer welfare.

A second alleged anticompetitive effect of the Visa and MasterCard issuance restrictions was that they limited the ability of American Express or Discover to issue offline debit cards. Offline debit cards are linked to the checking accounts of the cardholders’ banks and thus are typically issued by the cardholders’ banks.” Debit card usage has been growing more rapidly

---

“American Express was not completely precluded from issuing debit cards. Through its ownership of American Express Centurion Bank, a traditional bank with some $11 billion in assets, Amex...
than usage of credit and charge cards, and in the future debit functionality may be linked to credit card functions on a single smart card. Thus, the issuance restrictions may have limited an important source of network growth for American Express and Discover. To the extent that this caused their networks to be smaller than they would have been otherwise, consumer welfare would again be reduced.

Pro-Competitive Effects of Issuance Restrictions

There were also pro-competitive aspects of the Visa and MasterCard issuance restrictions. As open associations, the Visa and MasterCard networks were vulnerable to free-riding by some of its members. Such free-riding, which could damage the quality and growth of the network, was prevented by the issuance restrictions, which in effect created a loyalty rule that forced all members of the network to invest in and promote the venture.

To understand this, suppose there were no issuance restrictions, so that American Express could sign issuance agreements with banks that issue Visa and/or MasterCard. An important way in which American Express differentiates itself is by attracting cardholders who are high per-transaction spenders. In fact, this is how American Express is able to justify to merchants its relatively high merchant discount. Thus, it is likely that American Express would seek partnerships only with a select group of issuers, namely banks that have a large customer base with high average spending or that have upscale private clients. If the largest and most profitable issuers partnered with Amex, those banks would still benefit from the investments made by the Visa and MasterCard networks and their other member banks. However, those banks would have less incentive to innovate and develop features to promote the Visa or MasterCard brands or to support network initiatives to introduce new cards that would compete directly with their Amex offerings. The reason is that other competing banks cannot decide to issue Amex and thereby take advantage of the Amex products. In contrast, Visa and MasterCard are open networks, so any bank can issue their cards and benefit from new Visa and MasterCard products. If, for example, Visa developed a new corporate card, all issuers would have access to this product and not just a select few.

This last point is important. One might ask why, if it is harmful to allow Visa banks to issue Amex products, isn’t it also harmful to allow those banks to issue MasterCard products? The reason is that as an open network, MasterCard cannot choose to allow only a select group of banks to issue its products (and likewise for Visa). MasterCard and Visa can compete with offered online debit cards to its checking account holders. Amex could also offer PIN-based and signature-based debit functionality to customers of any bank if those customers authorized Amex to access their checking accounts. Nonetheless, the ability of Amex to expand its debit card base was clearly limited.
each other for increased issuance at more banks, but they cannot restrict their membership to a limited set of banks.

Preventing such opportunistic behavior—a few select banks partnering with American Express or some other closed network, at the expense of other banks in the network—was at the heart of the issuance restrictions. To the extent that this strengthened the Visa and MasterCard networks, there was a benefit to consumers.

The Court’s Opinion

In its decision, the court found that “defendants’ exclusionary rules restrict competition between networks and harm consumers by denying them innovative and varied products.” Thus, the court ordered Visa’s Bylaw 2.10(e) and MasterCard’s Competitive Programs Policy be repealed and enjoined “any further prohibition by the defendants of the issuers’ ability to issue general purpose and debit cards on other general purpose networks.” The decision noted that “the abolition and prospective injunction of defendants’ exclusionary rules will open the market to American Express and Discover to compete with MasterCard and Visa to enter into card issuing arrangements with banks. The combination of the distinct characteristics of the American Express and Discover networks with the specific attributes and issuing competencies of the issuing banks will result in increased output and consumer choice, in addition to strengthening the networks by increasing their scale and relevance.”

The court disagreed that removing the Visa and MasterCard issuance restrictions would lead to opportunistic behavior significant enough to weaken either of the two networks, and the court believed that any such harm would be outweighed by the benefits that would accrue to consumers. As the court stated in its decision: “Multiple bank issuance of general purpose cards strengthens general purpose credit and charge card networks in three fundamental areas: increased card issuance, increased merchant acceptance, and increased scale. When combined with new products and services that bank issuance provides—such as the practical ability to offer customers a debit product on the network infrastructure—strengthening the networks in these areas benefits consumers both directly (by ensuring

19 The Visa and MasterCard issuance restrictions did not apply in Europe, where banks that issued Visa and/or MasterCard could also issue Amex products and other local cards. One might examine what impact, if any, this had on the Visa and MasterCard networks in Europe. The European experience, however, received little attention at trial, in part because the number of issuers there is much more limited.

21 United States v. Visa USA Inc., at 408.
22 United States v. Visa USA Inc., at 408.
availability of new products and services) and indirectly (by lowering net-
work costs that are passed on to consumers).”’

**COMPETITIVE EFFECTS OF DUAL GOVERNANCE**

At the time of the Government’s lawsuit, duality had been in place for some twenty years and, to use a distinction made by the DOJ, existed at both the issuance and governance levels. **Issuance duality** refers to the fact that bank and nonbank issuers on one network could also issue on the other network. **Governance duality** refers to the fact that the boards and advisory committees of each network could have members from institutions that issue on the competing network, and a member institution could have representatives on the board of one association and on an advisory committee of the other association. The DOJ did not allege harm to competition from issuance duality by itself; its concern was with governance, but it argued that issuance duality was part and parcel of the problems that were created by dual governance.

The DOJ claimed that because of dual governance, MasterCard members did not promote the MasterCard brand at the expense of the Visa brand, and vice versa. More specifically, the DOJ argued that duality diminished the overall intensity of network competition between Visa and MasterCard, which resulted in reduced levels of advertising, innovation, and the development and introduction of new types of cards. Visa and MasterCard argued that issuance duality versus governance duality is a meaningless distinction because duality basically boils down to overlapping members. Visa and MasterCard also argued that duality increased the level of competition among networks for bank issuers and that there was no evidence of reduced levels of advertising innovation linked in any way to duality. Finally, as the smaller of the two networks, MasterCard argued that duality was essential in order to preserve the MasterCard network as an effective competitor.

**Anticompetitive Effects of Dual Governance**

Banks that issue (and thus exert control over) Visa have an interest in the success of the Visa network. Likewise, banks that issue (and exert control over) MasterCard have an interest in the success of the MasterCard network. For the most part, however, these have been the same banks. The concern, then, is that these dual members may have financial incentives to reduce investments aimed at shifting share between the two networks. Advertising is an example of this; although MasterCard’s advertising will to some extent take share from American Express and Discover and will

*UnitedStates v. Visa USA Inc.*, at 387.
also generate new cardholders, at least some of the gain will come from shifting Visa cardholders over to MasterCard. Thus, at least in theory, duality could lead to some reduction in advertising by MasterCard and Visa. Some advertising, of course, is merely "persuasive" and has no informational content. However, to the extent that "suppressed" advertising would have been informative, it could lead to some reduction in consumer welfare.

It is also possible that duality might lead to a reduction in investments aimed at technological innovation or the development and introduction of new types of cards, insofar as those investments would cause a shift in share between the Visa and MasterCard networks. Indeed, the DOJ claimed that the development of chip-based smart cards and investments in encryption technology for Internet transactions by Visa and MasterCard were hindered for just this reason. To the extent that reduced incentives to compete indeed delayed or reduced investments in innovative technologies of this sort, it would mean a reduction in consumer welfare.

Of course, it is impossible to know how much advertising and investments in new technology would have occurred in a "but-for" world without duality. But, at least in theory, without duality the managers of Visa and MasterCard might have had a stronger incentive to compete against each other, possibly by differentiating the two networks to a greater extent. And it is possible that greater differentiation (in whatever direction it went) would have created products that better satisfied the tastes of some consumers. To the extent that duality stifled such product differentiation, it could have caused a reduction in consumer welfare.

Pro-Competitive Effects of Dual Governance

There are also pro-competitive aspects to duality. One of the most important is that it helps to ensure the survival of MasterCard, whose network is considerably smaller than Visa's. If issuing banks could be a member of only one of the two networks, it is likely that most would choose Visa. The Visa network would then grow further at the expense of MasterCard, and a "tipping point" might be reached at which nearly all issuers abandoned MasterCard so that they could join the Visa network. We would then be left with just one of the two networks, which cannot possibly yield an improvement in consumer welfare. (Not surprisingly, MasterCard was especially concerned that the court might rule to end duality. And why, then, did Visa defend duality? Because, as discussed below, it was in the interest of its members.)

In addition, it is not clear that duality leads to a net decrease in the intensity of competition between the Visa and MasterCard networks. The reason is that duality encourages the two associations to compete directly for issuers by making it easier for banks to switch part or all of their portfolios from one network to the other. Member banks use duality to extract concessions from the networks and pass a portion of these incentive
payments on to consumers in the form of lower fees, lower interest rates, and increased rewards programs, thereby differentiating their cards from their competitors. For example, when banks consider new initiatives or the targeting of a new group of consumers, Visa and MasterCard typically approach them to promote their networks and cards and might offer incentive payments to support marketing or product testing in an effort to convince banks to go with one association over the other.

Finally, because issuers can (and many do) offer cards over both the Visa and MasterCard networks, they can benefit from initiatives taken by either association by changing the focus of their solicitations to consumers. This is analogous to a store’s ability to promote competing brands of a product. Without duality, issuers would be able to take advantage only of innovations developed by the network of which they are members. To take advantage of a competing network’s innovations, the issuer would have to convert its entire portfolio of cardholders over to that network. With duality, members of the associations benefit from the development, promotion and adoption of new card features and technologies because these enable members to compete with other issuers as well as with other card networks. This increases the intensity of competition between the two networks because one of the important ways in which they compete is through, and for, issuing banks.

The Court’s Opinion

In its decision, the court found that "the Government has failed to prove that the governance structures of the Visa and MasterCard associations have resulted in a significant adverse effect on competition or consumer welfare." The Government had asserted that duality would reduce competition between the two networks, and thereby lead to less advertising, and a reduction in investments in innovation and in the development and introduction of new types of cards. The court found, however, that the Government had failed "to establish causation between dual governance and any significant blunting of brand promotion or network and product innovations" and that such failure "is fatal to this claim."

In fact, the court went further. It noted that with respect to duality, "defendants had no obligation to demonstrate its pro-competitive effects. Nonetheless, the record evidence demonstrates that in some instances dual governance had pro-competitive effects, most notably by facilitating share-shifting competition by MasterCard." Thus, Visa and MasterCard have been free to retain duality—both in issuance and in governance.

*United States v. Visa USA Inc., at 328.
CONCLUSION

In summary, the court ruled that the Visa and MasterCard associations could retain duality but also ruled that their issuance restrictions were anticompetitive and had to be removed. Judge Jones's decision regarding the issuance restrictions was appealed by Visa and MasterCard but was affirmed by the Second Circuit on September 17, 2003. The associations then appealed to the Supreme Court, which denied certiori in October 2004.

What impact did the outcome of this case have on competition among payment card networks and on the operation and growth of Visa and MasterCard? Remarkably little. In the end, duality remains in place, and the two networks have not been forced to change their systems of governance. MasterCard has since become a publicly traded corporation, but the member banks control most of the stock, and duality and governance have not been materially affected.

On the other hand, Visa and MasterCard were forced to abandon their membership restrictions, so that banks issuing cards under either (or both) networks are now free to negotiate issuing agreements with American Express, Discover, or any other network. During the trial, Visa and MasterCard argued that this would lead to opportunistic behavior whereby some select banks would partner with, say, American Express at the expense of other banks in the network, thereby weakening the Visa and MasterCard networks. The Government argued (and the court agreed) that removing the membership restrictions would strengthen the American Express and Discover networks by allowing them to compete more aggressively with MasterCard and Visa, resulting in increased output, product innovation and differentiation, and thus lower costs and greater choice for consumers.

In fact, however, American Express and Discover have remained largely closed networks. By the end of 2006, American Express had signed several issuance agreements (the largest with MBNA/Bank of America and Citibank), but only about 5.1 million Amex cards, out of a total of 48.1 million, had been issued by third-party banks. Furthermore, the 2006 transaction volume from these bank-issued Amex cards was only $7.81 billion out of a total Amex volume of $406.80 billion.

Overall, the Visa and MasterCard networks do not appear to have suffered in any measurable way from American Express' and Discover's ability to issue through banks. Furthermore, although output (as measured by charge volume or the numbers of cardholders) has continued to increase,

---


28 Visa has announced its intention to become a publicly traded company as well.

and there has been ongoing innovation and differentiation, there is no indication that this picture would look any different had the membership restrictions remained in place. Put simply, Visa’s and MasterCard’s fears that many of their larger issuers would start issuing Amex cards have not come to pass. Why?

Recall that there are important advantages to being a closed network—for example, a better ability to control standards and service quality, control over cardholder data, and potentially a better ability to cross-sell other products and services. In addition, it is important for American Express to maintain the identity of its payment cards as geared to higher-income and higher-spending consumers and avoid cannibalization of its proprietary cards. If bank-issuers target consumers similar to Visa and MasterCard cardholders, American Express risks erosion of its premium brand; and if they target consumers similar to current Amex cardholders, there would be a greater risk of cannibalization. Thus, when negotiating issuance agreements, American Express will quite naturally be very selective in its choice of banks and very demanding in the terms of any agreement. The larger and stronger banks (the ones that American Express might be interested in) will likewise be very demanding, as they have no interest in cannibalizing their Visa or MasterCard portfolios, and will want up-front payments and an interchange fee sufficient to offset the startup and ongoing costs of issuing American Express. As a result, we would expect that issuance agreements will be difficult to reach, and that has indeed been the case.

Payment card networks are continuing to grow. Increasing numbers of merchants of all types (e.g., fast food outlets) have begun accepting payment cards, as they find that their cost effectiveness relative to cash or checks easily outweighs the merchant discount. As they gain wider merchant acceptance, and as the networks and their issuers continue to differentiate their card offerings, more consumers prefer using cards as a means of payment. All four networks are investing heavily in innovation, much of which is directed at what is currently one of the greatest problems for continued network growth—fraud and other threats to card security. Another limit to network growth is saturation; in the United States, most creditworthy consumers already hold several cards. As a result, the networks are investing heavily in international expansion, which has been progressing rapidly. The Government’s antitrust case has raised interesting issues but so far has had little impact on the industry.

REFERENCES


INTRODUCTION

In May 1998, the U.S. Department of Justice (DOJ), twenty individual states, and the District of Columbia filed suit against the Microsoft Corporation claiming that Microsoft had monopolized the market for personal computer (PC) operating systems and had used its monopoly to engage in a wide range of antitrust violations. The case was tried in federal district court from October 19, 1998, through June 24, 1999. The court reached its findings regarding the facts of the case on November 5, 1999, and its legal conclusions on April 3, 2000.

Microsoft’s appeal to the Circuit Court of Appeals for the District of Columbia was decided on June 28, 2001, followed by extensive settlement discussions among the various parties—the Department of Justice, the states, and Microsoft. The appellate court affirmed the monopolization claim, reversed other findings by the district court, and remanded the case back to the district court to find an appropriate remedy. Following further settlement discussions among the various parties—DOJ, the states, and Microsoft—DOJ and Microsoft reached a settlement agreement. Nine states opted not to join the settlement, proposing a different remedy. A thirty-two-day remedy trial was held, and, on November 1, 2002, the district court
issued a remedy ruling. Eventually, all of the remaining states settled with Microsoft.

While its ultimate impact on antitrust jurisprudence remains unclear, there is no doubt that from the public’s perspective U.S. v. Microsoft was the antitrust case of the 1990s, and arguably from a policy perspective one of the most significant antitrust cases of the twentieth century. The investigation, the trial, and its aftermath received wide press coverage throughout. A number of the major actors in the drama became household names, as much the result of the public relations battle among the parties as of the litigation itself.

Microsoft’s problems did not end with the resolution of the U.S. Government’s case. Microsoft was sued in class actions brought by attorneys on behalf of customers who were alleged to have paid too much money for their Windows operating systems. Most of these lawsuits have now been settled for very substantial sums of money, running into the hundreds of millions of dollars.

To make matters worse from Microsoft’s perspective, the turn of the century coincided roughly with increased enforcement activity by the Directorate General for Competition of the European Union. Rather than focusing on browsers and on PC operating systems, the EU directed its attention to the “player,” which allows users to stream audio and video content from the Web and to servers. Real Networks, which dominated the market with its Real Player, was being threatened by Microsoft’s decision to integrate the Microsoft Player into its operating system.

After an extensive investigation, the European Commission concluded in March 2004 that Microsoft’s bundling of its operating system with its player constituted a violation of the EC law that characterized certain anticompetitive monopolizing practices. The EC ordered Microsoft to pay a substantial fine and to put onto the market a second version of its operating system Windows XP without a player. Microsoft’s appeal of the EC’s ruling (to the Court of First Instance, an intermediate appellate court) led to a September 2007 ruling that strongly supported the EC. As of the early autumn of 2007, the EC continues to monitor and to question Microsoft’s responsiveness to its required remedies.

In sum, there is much in the Microsoft case that is of import to the twenty-first-century competition in high technology. Not only does Microsoft’s behavior continue to be scrutinized, but also courts are now beginning to look at the implications of the Court’s ruling in U.S. v. Microsoft on other high-technology companies such as Intel and Google.

At the heart of the Microsoft case was the Government’s claim that Microsoft had engaged in a range of anticompetitive acts that was designed to maintain its operating system (OS) monopoly. The Government did not question the source of Microsoft’s historical success. The Government did, however, claim that consumers were harmed by Microsoft’s conduct, in part because consumers were paying higher prices for their operating
system software and in part because Microsoft’s actions had reduced innovation in the software industry. In response, Microsoft argued that it was not a monopoly since it faced significant competitive threats in a highly dynamic industry. It further argued that its success should be seen as pro-competitive, since consumers had benefited as the result of the distribution of its high-quality, innovative software. If the court were to impose substantial antitrust remedies, Microsoft believed, competitive incentives would be reduced, which would lead to less, rather than more, innovation.

BACKGROUND

The Microsoft Corporation is a relatively young corporation, having begun its existence in the mid-1970s. Since that time, Microsoft has enjoyed exceptional and unparalleled growth and in the process has made millionaires of thousands of its employees. Many attribute this success to Microsoft’s skill and foresight, while others cite an element of luck; all would most likely agree, however, that Microsoft has shown an uncanny ability to adapt its business plans and to market innovative technology successfully. With success, however, has come antitrust scrutiny, with various governmental agencies and private plaintiffs questioning whether Microsoft has used a range of practices to restrain competition, to exclude competitors, and to expand its market power beyond the operating system market.

The first governmental investigation of Microsoft began with the Federal Trade Commission (FTC). In 1990 the FTC undertook an investigation of Microsoft’s software licensing practices with personal computer original equipment manufacturers (OEMs). After nearly three years of investigation, the five-member Commission failed to support the legal staff’s desire to bring a suit against Microsoft (the vote was 2-2; one commissioner had not participated).

This victory for Microsoft was short-lived; the Department of Justice undertook its own investigation almost immediately. A year later, on July 15, 1994, DOJ filed a complaint, claiming that Microsoft’s contracts with OEMs were exclusionary and anticompetitive and that their purpose was to allow Microsoft to maintain its monopoly in the market for personal computer (PC) operating systems (OSs). The case did not go to trial. Microsoft and the Government settled, with Microsoft signing a consent decree in which it agreed to restrict its licensing agreements along a number of dimensions. (The agreement was finally approved on June 16, 1995.) An important aspect of the consent decree was the agreement that Microsoft could not condition or “tie” its operating system license to the license of

---

2 See, for example, Gilbert (1999).
other operating system products. However, the agreement did explicitly allow Microsoft to continue to develop “integrated” products.

The distinction between an anticompetitive tie and pro-competitive product integration was to become a central issue in the Microsoft litigation that followed. But there was an additional skirmish to be fought before the larger battle began. With the rapid development of the Internet came the need for software that would allow PC users to have easy Internet access. The first highly successful web browsing product came from Netscape. In a very short period of time Netscape’s Navigator became the market leader, accounting for approximately 70 percent of browser usage in 1996.

While initially slow to realize the potential significance of the Internet, Microsoft was quick to redirect its efforts aggressively toward Internet browser software in 1996. A new antitrust issue arose when Microsoft required OEMs to license and to install Microsoft’s browser, Internet Explorer (IE), into new PCs as a condition for obtaining a license to install the Windows 95 operating system. The Government sued, claiming that Microsoft’s tie between IE and the OS violated the 1996 consent decree. Microsoft defended by claiming that IE and the OS were integrated products and that consequently its licensing arrangement should be seen as an exception to the 1995 consent decree. The Government was initially successful: On December 11, 1997, Judge Thomas Penfield Jackson ordered that Microsoft separate its Windows 95 OS and IE. However, on appeal the U.S. Court of Appeals for the District of Columbia sided with Microsoft, claiming that Microsoft had offered evidence that the combination of IE and the OS offered functionality that was not available without product “integration.” In a move that would later be seen as prescient, the appellate court made it clear that its decision was based solely on its reading of the consent decree and not on broader antitrust principles.

The battle began on May 18, 1998, when the U.S. Government, twenty of the states, and the District of Columbia brought suit against Microsoft. In its filing, the DOJ alleged that Microsoft had engaged in a range of practices involving operating system licenses with OEMs, contracts with Internet service providers (ISPs), and ties between the OS and its IE browser, all of which restrained trade in violation of Section 1 of the Sherman Act. The DOJ also alleged that Microsoft had attempted to monopolize the market for Internet browsers in violation of Section 2 of the Sherman Act. The heart of the Government’s case was its allegation that Microsoft had engaged in a range of practices that aimed primarily, if not solely, to protect and maintain the Microsoft operating system monopoly.

Some background will be useful at this point. In the years prior to the filing of the Government’s case, it became apparent to Microsoft that the

5 Unless otherwise stated, “Government” applies to the U.S. Department of Justice, the states, and the District of Columbia.
Netscape Navigator browser could serve as the foundation for a software “platform” that had the potential to compete with Microsoft’s Windows 95 (and later Windows 98) operating system. Operating systems provide application programming interfaces (APIs) through which applications interact with the OS and through the OS with the computer hardware. Applications developers must write their programs to interact with a particular operating system’s API. The time and expense of then “porting” the application to a different OS can be substantial. The term platform describes a set of APIs to which applications may be written.

Because of the huge volume of Windows operating system sales and the size of the network of OS users, a vast number of applications, including the highly successful Office suite, have been written for Windows. If a firm were successfully to offer a competing OS, it would, of necessity, need to offer a substantial number of applications, which would most likely include word processing and business productivity software. As Apple realized during the 1980s, failure to offer a range of applications that appeal to businesses is likely to hinder one’s ability to grow market share.6 Because much of the software development and marketing effort is sunk (i.e., cannot easily be of value elsewhere), the result is the presence of a significant "applications barrier to entry" to OS markets. The ability to reduce the significance of this barrier to entry is what made the Java programming language of particular interest to Microsoft and its competitors.

Netscape Navigator (and other browsers) relied on Java. Developed and marketed by Sun Microsystems, Java was a “cross-platform” language that offered to applications programmers the opportunity to write a program once but to have that program run on all operating systems. Cross-platform Java effectively served as a form of “middleware,” software that sits on top of an OS while at the same time serving as the foundation for other applications. According to the Government, Netscape threatened Microsoft because its browser had the potential to distribute cross-platform Java to independent software developers. If those developers chose to write to other OSs such as IBM’s OS/2 or Linux (or if they wrote directly to browser APIs associated with Internet applications), the Windows monopoly would be at risk.

With respect to the maintenance-of-monopoly claim, the Government alleged that Microsoft had engaged in a range of practices whose purpose was to limit severely the commercial viability of the Netscape browser. These included an attempted market division with Netscape and attempts to discourage Apple and Intel from participating in closely related, and potentially threatening, software markets. They also included a combination of exclusionary devices (including those mentioned previously) and

---

6 In response to this concern, Apple licensed Windows 1.0 to Microsoft in exchange for a Macintosh version of the spreadsheet program Excel (and a one-year delay before the Windows version of Excel was offered).
other predatory acts aimed in part at Java and in part at the Netscape browser, an important means by which Java can be distributed to computer users. Exclusionary behavior is behavior that makes it more difficult for competitors to compete; predatory behavior arises when a firm forgoes short-run profits to drive a competitor from the market in the hope of recouping future profits in the same market or in some other market.

In its April 5, 2000, opinion, the district court found in favor of the Government with respect to its Section 2 claims, including both its core maintenance-of-monopoly arguments and a separate attempted-browser-monopolization claim. While determining that the browser-operating system tie and certain contractual practices violated Section 1 of the Sherman Act, the court found that the Government had not shown sufficient evidence that the exclusionary practices had foreclosed competition.

Judge Jackson then accepted the Government’s proposed remedies, which included conduct remedies that would limit Microsoft’s use of exclusive contracts and its control over the PC “desktop.” He also supported the Government’s proposal that Microsoft be divided into two smaller companies—an operating system company and an applications company. The latter would maintain control over the browser business, although the OS company would retain property rights with respect to the current version of the browser.

Judge Jackson agreed to stay his remedies until after the appeals court heard the case. In a highly unusual move consistent with the importance of the case, the entire eligible membership of the Appeals Court for the District of Columbia (rather than the usual three-judge panel) heard Microsoft’s appeal of the case. The unanimous appellate decision was seen as a victory by the Government, but it did contain positive elements for both sides. From Microsoft’s perspective, the opinion was successful because the appeals court (1) reversed Judge Jackson’s decision that Microsoft had attempted to monopolize the browser market, and (2) remanded (sent back) the case to the lower court for a rehearing on the remedy issue. From the Government’s perspective, the decision was highly successful because (1) the court upheld both the maintenance of monopoly claim, and (2) along with it, the court condemned the wide range of exclusionary and predatory practices alleged by the Government.

**AN OVERVIEW OF THE ECONOMIC ISSUES**

Two significant economic questions raised by *U.S. v. Microsoft* were:

1. Did the Microsoft Corporation possess monopoly power in the market for personal computer operating systems?
2. Did Microsoft maintain its monopoly power by a series of anticompetitive actions that unreasonably restrained trade?
In answer to the first question, the Government stated that Microsoft did possess monopoly power in the market for operating systems for Intel-compatible desktop personal computers. Microsoft responded that the relevant market for antitrust purposes is substantially broader than Intel-based PC OSs; it includes hand-held computer operated systems and servers. Moreover, Microsoft faces substantial competition from other computer OSs, threats from other non-OS platforms that can support applications, and threats from yet unknown innovations. Indeed, the very fact that Microsoft found it necessary to take action against Netscape and Java shows that those companies and their products are in the market. Consequently, Microsoft does not have monopoly power.

With respect to the second question, the Government argued that Microsoft foresaw the possibility that the dominant position of its Windows OS would be eroded by Internet browsers and by cross-platform Java. Microsoft engaged in a series of anticompetitive practices in order to protect the monopoly power of its Windows OS. Microsoft responded that it did perceive a competitive threat from Java and responded in a number of ways to combat that competitive threat. However, those responses were the reasonable and appropriate responses of a competitor and cannot be appropriately characterized as an attempt by Microsoft to maintain its OS monopoly.

The Government raised an additional question that focused specifically on the market for browsers. That question raised a third issue:

3. In order to thwart the competitive advantage of the Netscape Navigator browser, did Microsoft use many of the same acts to hamper severely Netscape's browsing software?

Because questions 2 and 3 raised similar concerns about the alleged anticompetitive acts by Microsoft, it is not surprising that the economic analysis associated with question 3 is closely related to the analysis of 2. While question 3 was supported by the district court, it was rejected by the circuit court of appeals, which argued that the Government had not proved that there was a separate and distinct browser market or that there were substantial barriers to entry in the browser market. To keep the discussion focused, this case study will analyze only claims 1 and 2.

As mentioned, the core of the Government's case lay in its second claim—that Microsoft had used illegal anticompetitive practices to maintain its OS monopoly. According to the Government, Microsoft's conduct, which preserved and increased barriers to entry into the PC operating system market (and which distorted competition in the market for Internet browsers), included:

1. tying IE to the operating system (in effect requiring manufacturers to acquire Microsoft's Internet browser as a condition of acquiring Microsoft's Windows OS), thereby severely hampering Netscape's
browser and blunting the threat that software developers, in writing for a browser platform, would create software for a non-Microsoft OS;

2. excluding browser competitors from the most efficient channels of distribution (OEMs and ISPs), thereby requiring competitors to use more costly and less efficient channels;

3. imposing agreements requiring OEMs not to remove Microsoft's browser or to substitute an alternative browser;

4. imposing exclusionary agreements on ISPs, requiring them to boycott or disfavor Netscape and other browsers; these included agreements not to promote, distribute, use, or pay for Netscape's browser (or to do so only on less favored terms); and

5. giving its browser away for free ("committing" itself to do so "forever") and paying others to take its browser.

The following section elaborates on the two central economic questions raised by the case, in the process describing first the positions maintained by each of the two sides and then by the courts.

DEBATING THE ECONOMIC ISSUES

Did Microsoft Have Monopoly Power?

The Government's Perspective

The Government alleged that Microsoft possessed monopoly power in the market for operating systems for Intel-compatible desktop PCs. To support that claim, the Government provided market share data that showed that Microsoft's share of PC operating systems was very high and had remained stable over time; indeed, its worldwide share of shipments of Intel-based operating systems had been approximately 90 percent or more during the 1990s. Furthermore, numerous OEMs—the most important direct customers for PC operating systems—believed that they did not have any alternative to the acquisition and installation of Microsoft's Windows operating system.

The Government chose to exclude Apple's OS from the relevant market, based on evidence that there was very little substitutability between Apple and Windows. However, Apple's exclusion did not play a significant role in the Government's analysis of the case, primarily because of its relatively small market share.

Of greater potential significance were non-desktop devices such as the Palm Pilot, which could arguably substitute at least in part for the PC. To the extent that they could, Microsoft's monopoly power in PC operating systems would become less important. From the Government's perspective,
it was not credible that a 5 or 10 percent increase in the price of Windows above a competitive level would make a large number of users choose Palm Pilots rather than PCs. Moreover, even if non-desktop devices did at some point become serious substitutes for the PC, the Government argued that Microsoft’s monopoly power would only become less important; it would not disappear.

Note that in both the case of Apple and the case of hand-held devices, the crucial question (according to the Government) was not whether the hardware products involved competed with Intel-based PCs. That would matter if one were considering an alleged monopoly of PCs themselves. Rather, Microsoft’s alleged monopoly was over the market for operating systems for PCs; hence, an appropriate test was whether a 5 or 10 percent increase in the price of Windows above a competitive level would induce customers to change OSs or choose other hardware. This seems implausible, given that the price of Windows was only a small portion of the price of a PC at current prices and would be even smaller had the price of Windows been reduced.

While high market shares are an important indicia of monopoly power, the Government’s case had a broader foundation. In particular, there was direct deposition testimony from PC manufacturers—the most important direct customers for Windows—that a 5 or 10 percent price increase by Microsoft would not make them shift away from Windows. Further, Microsoft’s internal documents made clear that the company did not take into account the prices of other operating systems in setting its own prices.

Both parties agreed that OSs are characterized by network effects. Users want an OS that will permit them to run all the applications programs they want to use. As a result, developers tend to write applications for the most popular OS; and applications software written for a specific OS typically cannot run on a different OS without extensive and costly modifications or add-ons.

There are other network effects as well. For example, operating systems are complex; they exhibit network effects in part because firms are reluctant to reinvest in retraining workers and in part because using multiple OSs vastly increases technical support costs. This gives a firm an incentive to have the same OSs for all its own computers and the same OS that is widely used by other firms. Other network effects include the ease of exchanging files and the opportunity to learn from others.

As the result of economies of scale and network effects, the parties agreed that Microsoft’s operating system success has led to many more applications (more than seventy thousand) being written for its operating

7 Although a hypothetical “small but significant non-transitory increase in price” is a central part of the DOJ-FTC Merger Guidelines paradigm for market definition, its use for monopoly determination is controversial. See, for example, White (2000).

8 A similar argument applied to servers.
system than for any other. This reinforced and increased Microsoft’s market share, leading to still more applications being written for Windows than for other OSs, and so on. According to the Government, this positive feedback effect created an "applications barrier to entry" that made it difficult or impossible for rival OSs to gain more than a niche in the OS market. Indeed, the Government believed that Microsoft’s share and power were not likely to be eroded by new entry as long as the applications barrier to entry remained strong.

The Government also considered Microsoft’s argument that it faced significant competition from its own installed base of users. If so, Microsoft could not raise the price of new operating systems without fear that the installed base of users would choose not to upgrade their software. Indeed, Microsoft argued that it needed to continue to innovate for the same reason. While this installed-base argument relates to its pricing of upgrades, it does not apply to new computers, which are bought largely to take advantage of developments in hardware or software.

Microsoft’s Response

Microsoft denied that it has monopoly power; indeed, Microsoft claimed that the Government’s market definition was invalid. Stressing the network effects story, Microsoft argued that the appropriate focus should be on platforms, not on operating systems. Platform software offers standard-ized routines that allow software developers to avoid re-creating code to perform standard operations. While the Windows OS was a platform (as were other OSs), Microsoft argued that other non-operating system software, "middleware," could also serve as a platform. For example, Lotus Notes is a popular form of middleware that is popular among network users because of its e-mail and other functionality. Another significant platform is the World Wide Web, which consists of servers that make information, including applications software, available to the public communications network.

Microsoft argued that it competed vigorously to maintain its position as provider of the leading software platform. Indeed, Microsoft maintained that had it not continually improved its software, it would have been displaced by competing platforms. Whether threatened by known or unknown, actual or potential competition, Microsoft argued that any market power that it enjoyed was temporary and therefore not fairly characterized as monopoly power.

Microsoft also disputed the Government’s view as to the significance of the applications barrier to entry. According to Microsoft, if this was a significant barrier, then neither Netscape nor Java could ever successfully attract the allegiance of independent software developers. But, according to Microsoft, that was inconsistent with the reality that Netscape and Java were serious threats, despite their initial lack of applications. (The Government
responded, in turn, that it never claimed that the barrier was insurmountable; had it been, Microsoft would not have engaged in its anticompetitive behavior.)

In another defense against the claim that it had monopoly power, Microsoft relied on a standard static model for short-run monopoly pricing to argue that its Windows OS price was far below the price that a monopolist would charge. Assuming that there was monopoly power, Microsoft's economist calculated an elasticity of demand for Windows from information relating to the elasticity of demand for PCs, its OS market share, and the marginal cost of Windows (which was very low). Applying a "rule of thumb for pricing," Microsoft calculated that the short-run monopoly price for Windows was approximately $1800, well in excess of the actual price (around $60). Hence, the Government was wrong to characterize Microsoft as having monopoly power.

The debate surrounding the short-run profit-maximizing price of Windows was an extremely active one during the trial, perhaps because it had implications for the private lawsuits that were to follow on the Government's case and perhaps because it focused analytical attention on crucial short-run versus long-run and static versus dynamic distinctions. The Government objected strongly to Microsoft's claims, pointing out that since the marginal cost of Windows software is very low, the short-run profit-maximizing action for Microsoft is to price where the elasticity of demand that it faces is unity (when costs are zero, this maximizes revenue and profit). According to the Government, this is true whether or not Microsoft has monopoly power. According to the Government, the correct conclusion must be that something other than short-run profit maximization is happening, and in particular, Microsoft must be taking its OS monopoly profits in other ways (e.g., through higher priced applications). Moreover, according to the Government, Microsoft's pricing of its operating system (in particular its contractual prices to OEMs) is consistent with long-run profit maximization by a firm with monopoly power in a network market.

The issue of how a profit-maximizing firm should price in a dynamic network market remains a complex one. In a network industry it is in any dominant firm's interest to account in its pricing strategy for a host of factors that could lead, other things equal, to a lower price than one would expect from a simplistic short-run theory. These include: (1) the value of keeping and growing one's installed base, the source of the significant network effects; (2) the possibility of creating increased demand for complementary applications, which in turn provides an additional revenue source; (3) the need to discourage software pirating; and (4) the imposition of

See, for example, Pindyck and Rubinfeld (2001), pp. 333-334. According to the rule of thumb, the profit-maximizing price of a product is given by \( P = \frac{MC}{1 + 1/E_d} \), where \( P \) is the price, \( MC \) is the marginal cost, and \( E_d \) is the elasticity of demand facing the firm.
binding restrictions on its OEM customers as part of its anticompetitive campaign to preserve long-run monopoly profits.

The Court's Perspective

While Judge Jackson's opinion did not focus on the pricing issue, his Findings of Fact (in November 1999) supported the Government's position on all significant market definition and monopoly power issues. The court understood that, when defining a market in a monopoly case, it is appropriate to emphasize the constraints on an alleged monopolist's power with respect to buyers of operating systems, not the constraints relating to producers of complementary products. Seen in this light, Navigator and Java were complements to the OS that could facilitate the writing of applications that were also complements. They were not substitutes and therefore should not have been defined as part of the same market.

The court also agreed that there was a significant "applications barrier to entry." While Apple had 12,000 applications and OS/2 had 2500, neither could compete with Microsoft, which had more than 70,000 applications, one of which was its dominant business suite, Microsoft Office.

The appellate court chose not to overturn Judge Jackson's findings of fact on these issues, in effect affirming the district court's finding that Microsoft had monopoly power in the operating systems market. The academic and policy debate about the nature of dynamic competition, the measurement of market power, and the appropriate framework in which to evaluate the pricing of firms with market power will, however, continue beyond the scope of the Microsoft case.

Did Microsoft Maintain Its Operating System Monopoly by Thwarting the Threat Posed by Netscape's Browser?

The Government's Case

The Government argued that Microsoft engaged in a range of acts whose primary purpose was to protect its operating system monopoly. The Government pointed out that in May 1995 Microsoft's CEO Bill Gates had warned his top executives that the browser could "commoditize" the OS. His fear was that if Netscape's browser was successful, programmers would be easily induced to write software for competing operating systems, to the detriment of Microsoft. The key to maintaining its monopoly was to thwart the success of the Netscape Navigator browser.

The Government emphasized that before Microsoft began giving away its browser for free, ISPs and retailers had distributed browsers separately from OSs; as a result, there was clear evidence that there had been demand for OSs without browsers and for OSs with a choice of browsers. This supported the Government's view that a free browser was not simply a
competitive strategy to penetrate the browser market, and it also supported the Government's proposed remedy—that Microsoft allow OEMs to choose which browser to offer, or indeed, to offer no browser at all.

The Government argued that by bundling its browser with its OS and giving away its browser for "free," Microsoft prevented companies from successfully entering the browser market unless they successfully entered the OS market. The necessity of this "two-level" entry effectively increased the barriers to entry into the OS market and thereby protected Microsoft’s monopoly in operating systems. The Government believed that Microsoft recognized the threat from Netscape Navigator, because it was an Internet browser capable of supporting applications that were OS independent. By lessening reliance on the OS, the browser, while not performing all the traditional functions of an operating system, could have provided opportunities for competing OSs by reducing the applications barrier to entry. In sum, the Government placed great weight on the fact that Microsoft was concerned that browsers could ultimately develop into alternate platforms, thereby threatening its Windows OS.10

This threat was real, because the Navigator browser runs on many different operating systems, including Windows, the Apple Macintosh OS, and various versions of UNIX, including Linux. Netscape’s browsers contain their own set of APIs to which applications developers can write their applications. As a result, applications can be developed that will run on browsers regardless of the underlying OS. Similarly, browsers could have reduced the power of the OS monopoly by facilitating the expansion of network computing, in which users with "thin clients" use a network to access applications residing on a server, rather than hosting the application on the PC itself.

According to the Government, Microsoft recognized that it could protect its dominant position in the PC operating systems market by gaining and keeping a large share of the business in Internet browsers and by preventing any other browser from gaining a share sufficient to threaten Microsoft’s platform dominance or remain viable as a platform. Moreover, if Microsoft’s Internet Explorer browser were the dominant browser and Microsoft decided to support only Windows-based technology, developers would have little incentive to create applications that were not Windows-based.

In furtherance of its argument concerning browser shares, the Government showed that Netscape’s market share (70 percent in 1996) had declined significantly by 1997, to the benefit of Microsoft, and that this

10In April 1996, Bill Gates wrote that: "Netscape's strategy is to make Windows and the Apple Macintosh operating system all but irrelevant by building the browser into a full-featured operating system with information browsing. Over time Netscape will add memory management, file systems, security, scheduling, graphics and everything else in Windows that applications require. The company hopes that its browser will become a de facto platform for software development, ultimately replacing Windows as the mainstream set of software standards. In Netscape's plan, people will get rid of their existing PC and Mac applications in favor of new software that will evolve around the Netscape browser" (April 10, 1996, "The Internet PC," Plaintiff's Exhibit 336).
pattern was likely to continue into 1998 and 1999, as in fact it did. In the Government’s eyes, the market for browser usage had “tipped” in favor of Microsoft. With hindsight, the tipping did occur; Microsoft’s current browser usage share is more than 85 percent.  

In sum, the Government believed that a range of anticompetitive acts was taken by Microsoft to exclude competition in Internet browsers. These were acts that Microsoft would not have undertaken except to foreclose competition and to protect the applications barrier to entry. The specific actions included:

**Market Allocation:** The Government claimed that Microsoft engaged in a series of market allocation efforts (involving Netscape, Apple, and Intel) whose ultimate purpose was to minimize the competitive threat to its OS monopoly.

In June 1995, Microsoft had a business meeting with Netscape, the alleged purpose being to solicit this emerging competitor to engage in a market allocation scheme in which Microsoft would agree to let Netscape offer its browser without competition in the server market, while Microsoft would be given control of the PC browser market. If Netscape had agreed (it did not), Microsoft would have succeeded in eliminating its only serious browser competitor.

The Government believed that Microsoft also engaged in similar conduct with Intel. When Intel proposed offering certain platform-level software that conflicted with Microsoft’s platform plans, Microsoft threatened, among other things, to withhold support for Intel’s new generation of processors. Accordng to Intel’s chairman, Intel ultimately “caved” and withdrew the effort, at least under its own brand, explaining, “Introducing a Windows-based software initiative that Microsoft doesn’t support... well, life is too short for that.”

Finally, Microsoft had entered into an agreement with Apple that required Apple to make IE its default browser on all its Macintosh operating systems. According to the Government, that agreement forced Apple to place all competing browsers in a folder, thereby removing other browsers from the desktop. Microsoft also limited Apple’s ability to promote other browsers and tried to discourage Apple from developing its QuickTime streaming software, a platform threat to windows.

The Government viewed Microsoft’s conduct with respect to Netscape, Intel, and Apple as consistent with its efforts to prevent browsers from

---

1. As of June 2000, the IE browser usage share was 86 percent (Clark 2000).
2. Specifically, Microsoft attempted to convince Intel to agree not to engage in platform competition with Microsoft by developing its Native Signal Processing technology, which would have endowed microprocessors with enhanced video and audio capabilities. Because the NSP technology would have been available for non-Windows platforms, it could have presented a threat to Microsoft’s monopoly power.
becoming a threat to the applications barrier to entry. In each case (1) Microsoft was confronted with platform-level software to which applications programs could be written; (2) platform-level APIs threatened to erode the applications barrier to entry into PC operating systems by supporting applications programs that could be used with multiple operating systems; (3) Microsoft responded by attempting to get the supplier of the potential alternative platform-level software to agree to withdraw from offering it and to concentrate instead on products that did not offer platform potential; and (4) Microsoft was prepared to act to preclude the supplier of a potential platform-level software from succeeding in offering the platform, even if Microsoft’s actions were not otherwise sensible from a business perspective.

**Predatory Pricing:** According to the Government, once Microsoft recognized the potential threat posed by Netscape’s browser, Microsoft began devoting at least $100 million per year to develop IE and tens of millions of dollars a year on marketing and promotion. Despite the significant browser-related costs it was incurring, Microsoft distributed its browser at a negative price. The IE browser was not only given away free, but also companies were paid and given valuable concessions to accept, use, distribute, and promote IE. While a free browser might be a profitable “penetration pricing” strategy to grow its market, the Government cited Microsoft’s internal documents to show that Microsoft undertook its browser development not to make money from browsers but to prevent Netscape’s browser from facilitating competition with Microsoft’s monopoly OS. Indeed, Microsoft had referred to IE as a “no-revenue product,” at the same time that Microsoft emphasized the importance of the browser to Microsoft’s competitive position.

When it made its decision to supply IE without charge, Microsoft estimated that from 20 percent to 50 percent of Netscape’s revenues came from licensing its browser (such revenues amounted to nearly $200 million per year). Microsoft’s decision to price its own browser below cost was thus made when it knew that Netscape was charging for its browser and that Netscape depended on those revenues to continue to compete effectively. Indeed, Microsoft candidly described its pricing of its browser to Intel in an effort to convince Intel not to do business with Netscape, saying that Microsoft was “going to be distributing the browser for free” and that “this strategy would cut off Netscape’s air supply, keep them from gaining any revenue to reinvest in their business.”

A predatory pricing strategy is one in which the predator forgoes current profits in order to eliminate or cripple a competitor, with the expectation of recouping those forgone profits at some point in the future. In the

---

14 Government Trial Exhibit 39.
15 Steven McGeady August 10, 1998, Dep. Tr. 16-17.
Government’s view, Microsoft’s pricing strategy did not make sense as an ordinary business practice. It was giving away something (a “no-revenue product”) that it had spent a lot of money to develop and distribute (forgone profits) and something for which the leading competitor was charging. It was only when Microsoft’s gains from preserving and extending its monopoly (recoupment) were included that Microsoft’s conduct appeared to be profitable. From the Government’s perspective, the preservation of Microsoft’s OS monopoly alone would permit recoupment. Indeed, the Government introduced contemporaneous Microsoft documents showing that the company’s zero (or negative) price for its browser was not considered a way to earn competitive ancillary revenues.

**Bundling and OEM Restrictions:** Although IE was not “tied” or “bundled” with the retail version of Windows 95 when it was first released in the summer of 1995, Microsoft did bundle IE with Windows 95 in distributing Windows 95 to OEMs, and IE was bundled with all Windows 98 OSs that Microsoft distributed through retail or OEM channels. (In Windows 98, the browser was “integrated,” having been designed to share extensive code with the operating system.) According to the Government, Microsoft made the decision to bundle IE and Windows in one form or another even though there was demand for browsers separate from the demand for operating systems.

The Government also argued that Microsoft made its bundling decision not to achieve efficiencies, but to foreclose competition. The Government was not arguing that bundling per se was anticompetitive. Instead, because Microsoft did not give OEMs the option of taking Windows without the browser, it thus compelled those OEMs and users who wished otherwise to take IE nonetheless in order to get Windows. This foreclosure of competition arguably had an immediate harmful effect on consumers, whose choice of browsers was restricted. The harm was not simply to consumers who faced limited browser choice; other harms resulted from the unnecessarily cumbersome OS and from the limited options for those who preferred not to use a browser.

Microsoft also recognized that OEMs wanted the ability to develop their own desktops and to substitute Netscape’s browser for IE. The Government suggested that Microsoft, fearing this threat in 1996, imposed screen and start-up restrictions to prevent OEMs from developing their own first screen or positioning competing browsers more favorably than IE. From the Government’s perspective, had Microsoft simply viewed the browser

---

“In the traditional terminology of economics, bundling relates to situations in which firms sell packages of two or more products for which there may or may not be separate consumer demand. Tying, a more general term (although often used interchangeably with bundling by economists) applies to cases in which there are separate demands for two products, but the consumer must purchase one product in order to obtain another. With this terminology, it is appropriate to describe the browser as having been bundled with and tied to the OS.
as a complement to its OS, Microsoft would have had a clear incentive to support all browsers. A successful Netscape browser, for example, would increase the demand for the Windows OS. However, according to the Government, Microsoft saw the Netscape browser as a platform that would support substitute operating systems. Facing the threat of a potential substitute, Microsoft’s OEM restrictions were rational, indeed profit-maximizing, but at the same time, they were also anticompetitive.

With respect to OEMs, Microsoft required the distribution of IE and restricted the distribution of other browsers. The agreements required OEMs that wanted to preinstall Windows 95 or Windows 98 on their machines to preinstall Microsoft’s IE also. The agreements also limited the ability of OEMs to promote other browsers or to substitute other browsers for IE. Indeed, until changes were prompted by an early 1998 stipulation between Microsoft and the Department of Justice, the agreements typically required that licensees not modify or delete any of the product software. This prevented OEMs from removing any part of IE from the OS, including the visible means of user access to the IE software, such as the IE icon on the Windows desktop or the IE entry in the “Start” menu.

Licensees were not contractually restricted from loading other browsers on the desktop. However, most OEMs preferred to load only one browser to avoid user confusion and the resulting consumer support costs and to avoid increased testing costs. In addition, some OEMs viewed the desktop and/or disk space as scarce real estate and were generally reluctant to preinstall more than one software title in each functional category.

Microsoft’s restrictions on the startup screen were modified just before trial so that OEMs had somewhat more flexibility than when the restrictions were imposed. However, IE was still required to be installed on every PC, and the IE icon could not be removed. The result, according to the Government, was a significant exclusionary effect that ensured that IE was the only browser on most PCs shipped by OEMs. By January 1999, Navigator was on the desktop of only a very small percentage of PCs.

Exclusionary Agreements with Internet Service Providers: According to the Government, Microsoft required the promotion and distribution of IE and restricted the promotion and distribution of other browsers by striking deals with Internet service providers in order to protect Microsoft’s business in operating systems. After OEMs, ISPs are the largest distributors of browsers.

While Microsoft’s agreements with ISPs allowed them to distribute other browsers, Microsoft’s contracts typically required that the ISPs not distribute other browsers to more than a relatively small fraction of their customers. Some ISPs had agreements that allowed them to distribute IE and Netscape without preferences; Microsoft’s documents used the term “IE Parity” to identify these companies.

Page 546
Microsoft also created a separate desktop folder for certain favored ISPs and entered into agreements with AOL, CompuServe, Prodigy, and AT&T to appear in it. The Government offered evidence that in doing so, Microsoft extracted promises from the ISPs not to deal with Netscape or to do so only on very unfavorable terms. Of particular importance was the agreement that Microsoft reached with AOL, which by early 1996 was being installed on a large number of PCs, to ship IE.

Typically, the ISP restrictive provisions involved percentage restrictions on shipping for larger ISPs and restrictions on promotional efforts for smaller ISPs. These limitations included: (1) requirements that 75 percent or more of the ISP software shipments include IE as the only browser and that the ISP not ship a competing browser unless a customer specifically requested it; and (2) limitations that restricted the total shipments of non-Microsoft browsers by ISPs.

The Government stressed that, in its agreements, Microsoft offered ISPs valuable space on its desktop as well as direct payments in the form of rebates or bounties. In exchange, Microsoft placed requirements on ISPs that hindered their ability to promote or distribute Netscape Navigator. The Government viewed these provisions as anticompetitive. Their purpose and effect were to reduce the ability of competing browser manufacturers to distribute and promote their browsers through leading ISPs.

**Microsoft's Response**

Microsoft agreed that the Internet presented both an opportunity and a threat. The Internet expanded the market for platform software generally, and the demand for Windows in particular. However, the Internet was also a threat to Microsoft because it offered competitors an opportunity to gain control of the communications and language standards for the Internet, which would leave Microsoft at a severe competitive disadvantage.

Microsoft characterized its browser battle as part of a larger effort to maintain its position as the leading PC provider of the Microsoft Windows software "platform," the software code that can be accessed through APIs, and which therefore offers a wide variety of features and services to software developers. Microsoft argued that its responses to this threat were appropriate responses by a competitor. Microsoft did believe that Netscape's Navigator could become a competing software platform if it added sufficient APIs to become attractive to software developers. If successful as a platform, Navigator could influence Internet standards in ways that would be adverse to Microsoft. Microsoft argued that, as a result, it had every right to compete aggressively with its most serious platform competitor.

Microsoft further argued that its actions against Netscape were not predatory because they were directed toward improving its own browser and OS and were not conditioned on crippling or eliminating Netscape as a browser competitor. Microsoft claimed that it had decided as early as 1993...
to incorporate Internet features into its Windows platform. It invested $100 million annually to develop new and improved versions of IE. Indeed, because of the success of those investments, the IE browser eventually equaled if not surpassed the Netscape browser in quality. However, because its Internet development did not proceed sufficiently quickly, Microsoft chose not to include the browser in Windows 95. Microsoft suggested that it had always planned to "integrate" its browser in its OS and that there were significant efficiencies associated with doing so. Its reward from investing in the browser was similar to the reward that it expected from other features that were integrated into the operating system; it came from the increased value of a Windows OS that would continue to remain the leading software platform.

With respect to the Government’s market allocation claims, Microsoft responded that it had legitimate business reasons for meeting with Netscape, Apple, and Intel. In each case, there were pro-competitive reasons to have joint discussions. Indeed, as a general matter, discussions concerning software interfaces and complementary products with other software firms can be valuable to both parties and to consumers. Microsoft emphasized the complementary relationship between the Netscape browser and its OS, and similar arguments were made concerning Intel’s software tools for video streaming, Apple’s QuickTime software, and Microsoft’s OS. Finally, in each case, Microsoft argued that no agreements were reached, that it did not intimidate its competitors into altering their behavior, and that its competitors continued to be successful. Thus, according to Microsoft, Apple continued to develop its QuickTime for Windows product, and Intel’s software development was delayed, but not stopped entirely.

With respect to pricing, Microsoft responded that firms frequently give away software for free or at a very low price in order to compete in what could be a “winner-take-all” market battle. Moreover, Microsoft argued that offering a free browser with Windows 95 was the natural step toward a goal of integrating valued features into its OS, as it did with its Internet browsing functionality in Windows 98. Microsoft additionally argued that Netscape had alternative sources of revenue, since it could sell advertising on its web portal, Netcenter. Finally, Microsoft argued that the Government had offered no compelling evidence that Microsoft’s policy of integrating features into its OS was not profit-maximizing apart from any predatory motives.

With respect to bundling and OEM restrictions, Microsoft believed that it was justified in restricting OEMs from altering the start-up process to preserve the quality and speed of the start-up process and to give each user a consistent, “seamless” experience. Microsoft also argued that it did not have substantial market power. Absent monopoly power (or, at least, significant market power), bundling is not a rational anticompetitive strategy. Rather, it is likely to be harmless and to serve a legitimate business purpose. Indeed, from Microsoft’s perspective, the benefit gained by
creating interdependencies between IE and Windows was greater than any anticompetitive effects of bundling that might be imagined.

Microsoft further claimed that there were efficiencies associated with bundling that would be lost were Microsoft forced to separate the browser functionality from that of the OS. Microsoft further argued that the Government's claim that every product design, particularly in the area of computer software, can make a plausible claim for some efficiency or benefit, missed the point. In Microsoft's view, browser functionality was an essential component of the operating system, and separating the two products once they have been combined would be impossible.

Finally, Microsoft claimed that the absence of IE would undermine the quality of the OS, to the detriment of users. Microsoft distinguished Windows 98 from Windows 95 and thereby dismissed as irrelevant the Government's view that Microsoft had provided ways to remove IE in Windows 95—a function that would most likely not have been provided if it led to a decrease in the quality of the OS. They also argued that the Government's showing that it is possible within Windows 98 to remove the ability to browse the Web with IE and to replace IE with another browser with no appreciable decline in the quality of the Windows 98 OS was misleading, since the Government's expert removed only the visible means of accessing the browser and not the browser functionality itself.

With respect to its ISP agreements, Microsoft took the position that it was justified in competing aggressively for the distribution of its browser. Its success with AOL, for example, was due to its superior quality product and the desirable competitive terms that it offered. Microsoft stressed that its browser came in distinct "modules," which allowed AOL more easily to integrate the IE "technologies" into its own proprietary software. Finally, Microsoft argued that Netscape did not offer the same support as Microsoft, refusing, for example, to allow AOL to manage its popular web portal.

With regards to both OEM and ISP agreements, Microsoft argued "no harm, no foul." Microsoft offered evidence to support the view that Netscape had many channels through which it could distribute its browser. It could "carpet bomb" by mailing disks with CDs, or it could encourage individuals to download IE from the Web. Consequently, neither Netscape nor any other competitor was foreclosed from competing with Microsoft.

The Court's View

Judge Jackson strongly supported the Government's claims that Microsoft had used anticompetitive acts to maintain its OS monopoly. The court's condemnation of Microsoft's behavior with respect to maintenance of monopoly was widespread. In each of the four categories of activities listed above—market allocation; predatory pricing; bundling; and exclusionary agreements—the court sided with the Government.
Moreover, with respect to the bundling arguments, the court agreed that “tying” the browser to the OS was anticompetitive even apart from its contribution to Microsoft’s maintenance of monopoly goal. Microsoft’s only “victory” at the district court level was Judge Jackson’s finding that the Government had not provided sufficient evidence that the OEM and ISP restrictions had foreclosed competition.

The circuit court of appeals supported Judge Jackson with respect to the maintenance-of-monopoly arguments, thus strongly backing the core of the Government’s case. Specifically, the court of appeals affirmed the district court’s finding that the OEM and ISP restrictions described previously were anticompetitive. The appellate court also agreed with the core of the lower court’s finding that the means by which Microsoft tied its browser to its OS was anticompetitive. Moreover, the appellate court made it clear that certain of Microsoft’s marketing allocation efforts with respect to Apple and Intel were anticompetitive. The appellate court did, however, raise questions about the bundling-tying issue itself, remanding this issue back to the district court for further analysis of the benefits and the costs associated with Microsoft’s bundling of IE and its OS.

Was Microsoft’s Alleged Anticompetitive Behavior Harmful to Competition?

The Government’s View

In the Government’s view, Microsoft succeeded in effectively excluding Netscape almost completely from the personal computer OEM distribution channel. OEMs that license Windows were required to take (and not remove) IE. For most OEMs, including the largest, that meant including only IE with the PCs they shipped.

In evaluating the effectiveness of Microsoft’s actions on browser competition, the Government argued that the relevant measure was Microsoft’s share of browser usage. The evidence of Microsoft’s foreclosure of Netscape and other browser competitors could be seen by comparing Microsoft’s share of browsers distributed by ISPs that made IE their default browser with ISPs that did not make IE their default browser. At the end of 1997 Microsoft enjoyed a 94 percent weighted average share of browser shipments by ISPs that agreed to make IE their default browser, compared with a 14 percent weighted average share of browser shipments by ISPs that did not make IE their default browser. Further, Microsoft’s weighted average share of browser usage by subscribers to ISPs who made IE their default browser was over 60 percent. In contrast, Microsoft’s

*Here, the judge arguably failed to note that the imposition of significant costs on competitors can harm consumers. Even if those rivals are not driven from the market, their ability to compete effectively can be reduced if their costs of browser distribution are raised significantly.*
weighted average share of browser usage by subscribers to ISPs who did not make IE their default was less than 20 percent.

Figure 20-1 shows Microsoft’s monthly share of browser usage by three categories of ISPs, from January 1997 through August 1998. The top line shows Microsoft’s share of usage among subscribers to AOL and CompuServe rising sharply. These companies (now merged) were chosen because they represented the largest ISPs (with a total of more than 11.5 million subscribers and about 65 percent of all subscribers to the services in the “Top 80” as of year-end 1997), and because AOL and CompuServe, as online service providers, were contractually restricted in their promotion and distribution of non-IE browsers to a greater extent than were most other ISPs.

The middle line shows Microsoft’s share for all ISPs. The bottom line shows Microsoft’s share for the ISPs within the “Top 80,” which Microsoft

![Figure 20-1 Microsoft's Share of the Browser Market](image-url)

**FIGURE 20-1** Microsoft's Share of the Browser Market

*Three-Month Moving Average of Usage by ISP Category*

Source: AdKnowledge, Inc.
listed as having "IE Parity" (ISPs whose browser choice was not known to be contractually restricted), which had ten thousand or more subscribers, and for which data were available.

From the Government's point of view, the effects were striking. Microsoft's share of "IE Parity" browser usage—the category that is contractually neutral—rose in twenty months from 20 percent to just under 30 percent. This increase included the effects of technological improvement in IE as well as the effects of Microsoft's bundling and tying. By contrast, the "All ISPs" line showed an increase in Microsoft's share from 20 percent to 49 percent. Finally, for AOL and CompuServe, Microsoft's share rose from just over 20 percent to over 87 percent. (It is worth noting that the dramatic jump in that share occurred before the introduction of IE4—the improved version of IE—in October 1997.) Thus, contract restriction—not browser quality or ISP choice—was largely responsible for the overall decline in Netscape share.

The exclusion of Netscape and other competing browsers from the OEM channel has been even greater. Although several OEMs sought to replace IE with Netscape, none was permitted to do so. And the fact that IE was required to be included meant in most cases that only IE would be included.

Recall that because of its innovations and success in creating and distributing the world's first widely used browser, Netscape initially had a very large share of the browser market. Microsoft's browser share at the beginning of calendar year 1997 was approximately 20 percent and had been significantly lower earlier. Regardless of how share is measured, it is clear that Microsoft's browser share increased dramatically, and Netscape's browser share fell sharply, over the years 1997 and 1998. Indeed, as mentioned previously, Microsoft's browser share continued to increase through 1999 as well, reaching over 85 percent by mid-2000. The Government's claim that Microsoft's goal wasn't necessarily to drive Netscape completely out of business, but only to prevent Netscape from becoming the browser of choice for most consumers, was borne out by history.

**Microsoft's Response**

Microsoft responded at trial that its conduct was profitable without considering any gains from reduced competition because the wide distribution of its browser caused more people to buy PCs to browse the Internet, with the result that Microsoft was able to sell more copies of its Windows OS. In other words, browsers can be complements to operating systems to the extent that the sale of browsers can be used to increase demand for Windows. Moreover, Microsoft argued that it expected to benefit by improving the quality of its browser through innovative product development.

Microsoft argued further that, while it did compete aggressively to defend its market position, none of the actions to which the Government
objected had harmed consumers, nor would they do so in the future. Moreover, Microsoft argued that it had not succeeded in eliminating the competitive threat posed by Netscape and Java. It argued further that AOL, which acquired Netscape during the trial, could choose to distribute the Netscape browser if it wished to do so. If AOL had decided to follow this avenue, the share of the browser market controlled by Microsoft would have diminished sharply.

Microsoft disputed the Government’s measure of browser market share, offering numbers that suggested that the market for browsers had not yet tipped in Microsoft’s favor. Microsoft further argued that the allegedly restrictive distribution agreements with ISPs did not foreclose Netscape, because Netscape had many opportunities to distribute its browser—through PCs and ISPs and through downloads. Microsoft further argued that Netscape could have marketed its browser more effectively but had failed to do so.

The real source of Netscape’s decline, according to Microsoft, was that Microsoft had won the battle of Internet browsing technologies on the merits—it simply had a better product. Rather than being anticompetitive, its actions to (1) invest $100 million annually in its browser; (2) invest heavily in the distribution of IE; (3) integrate its browser into Windows; and (4) reach a contract with AOL to have AOL use IE technologies for its own subscribers, were pro-competitive. From Microsoft’s perspective the Government failed to distinguish these pro-competitive acts from a narrower set of acts that the Government viewed as anticompetitive.

The Court’s Perspective

Judge Jackson’s Findings of Fact supported the Government’s view that Microsoft’s anticompetitive acts caused immediate harm. According to the court:

To the detriment of consumers, . . . Microsoft also engaged in a concerted series of actions designed to protect the applications barrier to entry, and hence its monopoly power, from a variety of middleware threats, including Netscape’s Web browser and Sun’s implementation of Java. Many of these actions have harmed consumers in ways that are immediate and easily discernible. They have also caused less direct, but nevertheless serious and far-reaching, consumer harm by distorting competition. (paragraph 409)

By refusing to offer those OEMs who requested it a version of Windows without Web browsing software, and by preventing OEMs from removing IE—or even the most obvious means of invoking it—prior to shipment, Microsoft forced OEMs to ignore consumer demand for a browserless version of Windows. . . . Those Windows purchasers who did not want browsing software . . . had to . . . content themselves with a PC system that ran slower and provided less available memory than if the newest version of Windows came without browsing software. By taking the
actions listed above... Microsoft forced those consumers who otherwise would have elected Navigator as their browser to either pay a substantial price (in the forms of downloading, installation, confusion, degraded system performance, and diminished memory capacity) or content themselves with IE. None of these actions had pro-competitive justifications, (paragraph 410)

Many of the tactics that Microsoft has employed have also harmed consumers indirectly by unjustifiably distorting competition. The actions that Microsoft took against Navigator hobbled a form of innovation that had shown the potential to depress the applications barrier to entry sufficiently to enable other firms to compete effectively against Microsoft in the market for Intel-compatible PC operating systems. That competition would have conducted to consumer choice and nurtured innovation... It is clear... that Microsoft has retarded, and perhaps altogether extinguished, the process by which... middleware technologies could have facilitated the introduction of competition into an important market, (paragraph 411)

Most harmful of all is the message that Microsoft's actions have conveyed to every enterprise with the potential to innovate in the computer industry... Microsoft's past success in hurting such companies and stifling innovation deters investment in technologies and businesses that exhibit the potential to threaten Microsoft. The ultimate result is that some innovations that would truly benefit consumers never occur for the sole reason that they do not coincide with Microsoft's self-interest, (paragraph 412)

RESOLUTION OF THE CASE?

The U. S. v. Microsoft case has now been resolved, and a remedy chosen. The path to a remedy has been circuitous. The appellate court made clear its distaste for a structural remedy that would break up Microsoft into two separate companies—an OS company and an applications company. The case was remanded to the district court, where Judge Colleen Kollar-Kotelly presided over settlement discussions. The U.S. Government and nine of the eighteen states that remained as plaintiffs reached a tentative settlement with Microsoft in which Microsoft would consent to a range of behavioral remedies. However, nine states and the District of Columbia opted not to join the settlement; they objected that the remedies were likely to be ineffective and pressed for stronger remedies. The court held a remedies hearing in which a broad range of issues was debated.

While emphasizing Microsoft’s attempts to thwart competition from Netscape, the Government had argued at the original trial, and continued to argue in its settlement discussions, that there is a broader issue—that Microsoft engaged in anticompetitive acts with the goal of stemming...
competition from middleware products that threatened its OS monopoly. Seen from this perspective, the issue of what is an appropriate remedy to restore competition and to achieve adequate deterrence remained an open issue at the remedy hearing.

In its original remedy argument posttrial, the Government took the position that behavioral remedies could serve a useful temporary role but that such remedies would likely be difficult and costly to enforce and could inadequately deter Microsoft's wrongful behavior. Structural remedies, on the other hand, are less regulatory and therefore less subject to extensive intervention by interested parties. The proposed breakup would have divided the company along lines that some would argue are inefficient and would not by itself guarantee increased operating system competition.

With the decision of the appellate court discouraging structural remedies, and its own skeptical concerns, the Department of Justice (now under the Bush administration) and the nine settling states chose to focus entirely on conduct remedies, and, contrary to the argument just outlined, took structural remedies off the table. Their proposed settlement contained three components. First, it attempted to prohibit Microsoft from foreclosing the OEM channel of distribution by eliminating restrictive licensing agreements and outlawing retaliatory measures against OEMs by Microsoft. Second, it attempted to keep the ISP distribution channel open by placing limits on Microsoft's ability to discourage others from developing, promoting, or distributing non-Microsoft middleware products. Third, the settlement offered a series of compliance measures whose goal is to enforce the terms of the settlement agreement.

Those states opposing the proposed settlement argued that the proposed behavioral remedy will be largely ineffective. Their primary concern was that the proposed settlement did not prohibit Microsoft from illegally bundling Microsoft middleware into the Windows OS. Absent such a remedy, there is nothing to limit Microsoft from technologically tying non-browser middleware software to the OS, when such software is seen as a potential platform that would compete with Windows. The opposing states also argued that the proposed consent decree will not effectively prohibit retaliatory conduct and restrictive licensing practices, and it will not effectively open the ISP channel of distribution. They also claimed that the proposed settlement would allow Microsoft to withhold vital technical information from developers of rival middleware. Finally, they argued that the proposed enforcement mechanism will be ineffective.

Judge Kollar-Kotelly's ruling was generally supportive of the settlement agreement reached between the DOJ and Microsoft. While the court rejected many of the more aggressive remedies proposed by the nine litigating states, the court did offer more aggressive, and potentially more effective, compliance procedures that were sympathetic to issues raised by
the litigating states. In a summary of its full opinion, the court suggested that its remedy "is carefully tailored to fit the wrong creating the occasion for the remedy . . . and is forward-looking in the parameters of relief provided. . . . [and] is crafted to foster competition in the monopolized market." Only the future will provide an answer as to whether a settlement that includes a complex set of behavioral remedies will indeed restore competition and effectively deter wrongly anticompetitive conduct.

It has now been a decade since the Government filed its largely successful case against Microsoft. It is interesting to reflect briefly on what has come to pass in the interim. For most of this decade, the Government's predictions about browser competition turned out to be valid. IE's share of the browser market grew to well over 90 percent in the first five years after the conclusion of the case, and its operating system monopoly (again with over a 90 percent market share) has continued almost unabated. Arguments that Microsoft's monopoly was a fleeting one were clearly off the mark.

However, computer software and hardware markets are highly dynamic and subject to change. A number of significant events suggest that the next decade may bring new challenges to both Microsoft and the competition authorities. First, new browsers such as Mozilla Firefox and Opera have begun to make inroads into Microsoft's browser dominance. Second, the increasing popularity of the Web has begun to shift power toward Web-based companies such as Google and Yahoo! Third, although still problematic, Linux-based operating systems are finally beginning to gain some traction in the PC market. Whether the new Microsoft Vista operating system will have the success of its prior counterparts remains an open question at this time. What is clear is that there will continue to be a role for active government review of competitive practices in software and related high-technology markets.

REFERENCES


CASE 21
Links between Markets and Aftermarkets: Kodak (1997)

Jeffrey K. MacKie-Mason and John Metzler

The difficulty of obtaining parts, technical information and diagnostic software has effectively kept 3rd party service suppliers out of the advanced equipment service market.

Both IBM and Xerox will sell spare parts, but we do not. This makes it more difficult for a third party to service our copiers.

INTRODUCTION

In 1987 seventeen small companies filed an antitrust lawsuit against the Eastman Kodak Corporation ("Kodak"). These companies, several of them literally "mom and pop" operations, had been trying to compete with Kodak for contracts to provide maintenance service to end-use customers who owned expensive, durable Kodak photocopier or micrographics equipment. Eleven years later, there had been two district court opinions, two from the Ninth Circuit Court of Appeals, and one from the Supreme Court. Kodak appealed its guilty verdict to the Supreme Court one last time, but the Court refused to hear the appeal, and the case finally ended in 1998. Entire conferences have been devoted to the antitrust economics issues raised by Kodak, and numerous articles have been published in both legal and

*MacKie-Mason was expert economist for the plaintiffs on liability issues and testified at trial. Metzler assisted in the economic analysis.

'Trial Exhibit (hereafter, "Ex.") 99 at 6762, a Kodak report on micrographic service.

Ex. 649 at 1315-1316, a Kodak internal document.
economics scholarly journals. Since the initial Supreme Court opinion in *Kodak*, there have been many closely related appeals court opinions, and they stand in sharply divided conflict on at least two issues. *Kodak* is one of the most significant antitrust cases of the last decade or two. It is also one of the most controversial, and the controversy is far from resolved.

In this case study we report on the history and status of *Kodak*. We focus on the economic issues. Although there is still sharp disagreement on the antitrust policy that should be followed in response to the economic issues, there is now a fairly broad agreement about the structure, assumptions, and results of the basic economic theories; remaining disagreements are largely about the facts specific to *Kodak*’s situation. After summarizing the relevant factual and procedural background, we describe the theories. We then describe the economic analyses presented by the parties in court. We close by discussing several issues in the law and economics of antitrust that *Kodak* raises but that have implications far beyond cases about durable equipment and maintenance.

**MARKET BACKGROUND**

This case concerned *Kodak*’s practices relating to parts for, and maintenance service on, micrographic equipment and high-volume copiers. These machines are "durable" goods: goods that are purchased with the expectation of gaining utility from them for an extended period of time. ("Consumption" goods are purchased for the one-time utility that they provide.) Many durable goods, perhaps most, require ongoing maintenance for continued utility. For example, a copier with a burned-out light bulb generates little utility beyond that of an ordinary countertop.

Specialized terms have been developed to describe the markets in which these goods are sold. Consider a car, as an example: Because the utility derived from the car is what brings consumers into these markets, the market in which the car is sold is the "foremarket" or "primary" market. Demand for maintenance is composed entirely of consumers of the primary market good: Consumers don’t demand brake service if they don’t own a car. Therefore, we refer to the market for maintenance goods as the "aftermarket."

---

3 The first three subsections summarize a large number of facts. We believe it is fair to represent these as the undisputed facts, though, of course, almost no presentation of facts in a case is completely undisputed. In the interest of space, however, we do not provide specific citations to the trial evidence to support most of these background facts. When the trial evidence is summarized, we provide citations.

4 Fore- and aftermarkets do not exist only for equipment and maintenance. For example, we could have a computer operating systems foremarket and application program aftermarkets.
Micrographics
Micrographics equipment is used for creating, filing, retrieving, viewing, and printing microforms. Microforms include microfiche, microfilm (for example, with back issues of *The New York Times*), and, for secret agents, microdots. These are variations on the same basic idea: Using optics and fine-grained film, it is possible to expose an image of printed matter onto film at a greatly reduced size, thereby reducing storage space. With another set of optics, the film image can be enlarged to the original size, then viewed or printed.

Kodak invented this process in the late 1920s. Over the succeeding seventy years, micrographics equipment developed to encompass cameras, film processors, film duplicators, readers (enlarge microfilm for viewing or printing), COM (computer output to microfilm, devices that “print” to microfilm rather than paper), and CAR (computer automated retrieval, largely software/hardware combinations to automate handling large libraries of microfilm). Kodak remained a significant manufacturer and seller in most of these equipment categories.

Photocopiers
Kodak began manufacturing and selling high-volume photocopiers in the mid-1970s. These are large machines, weighing hundreds of pounds and generally selling for tens of thousands of dollars. High-volume photocopiers can handle from sixty thousand to more than one million copies per month.

Maintenance
Both copiers and micrographic equipment require extensive, ongoing maintenance, consisting of service labor and parts. In this case, essentially all maintenance calls included a service component and many, perhaps most, required parts as well. Labor and parts are not required in fixed proportions. Among other things, they can vary due to the choice between part repair and replacement and the use of preventative maintenance.

Kodak established a national network of service technicians to provide maintenance on its copiers and another technician network for maintenance of its micrographics equipment. Kodak advertised the quality of its maintenance.

By the early 1980s, there were many, small independent service organizations (ISOs) providing maintenance on Kodak micrographic equipment. There were a few ISOs servicing Kodak photocopiers. These ISOs typically provided maintenance at prices 15 to 30 percent below Kodak.

---

*We use the present tense to discuss the technologies in this case, but we are referring to technologies current as of the issues in the case, covering approximately the years 1980-1995.*
They provided service generally of the same quality as Kodak's. They at times provided customized maintenance options that Kodak would not provide.

The "Parts Policy"

In 1995 Image Technical Services (ITS) had won a large contract from the Computer Service Corporation (CSC) for micrographics maintenance. Kodak had been servicing CSC's equipment for about $200,000 per year with a four-hour guaranteed response time. ITS submitted a bid of $150,000 with the same four-hour response time guarantee. Kodak countered by lowering its bid $135,000. ITS came back with a bid of $100,000 and agreed to put a service technician on-site full time.

Shortly thereafter, Kodak instituted a policy of no longer selling parts for either copier or newly introduced micrographics equipment to ISOs. This was clearly a change in practice for micrographics. Kodak had previously sold parts to anyone who ordered them and seemed to support the small ISOs. Kodak agreed that this was a change in policy but claimed that the policy applied only prospectively: Kodak would refuse to sell parts only for equipment models introduced after the policy was announced.

It is not clear whether the change in practice on copier parts sales was also a change in policy. Kodak had previously sold copier parts to some ISOs and had referred the ISOs to Kodak's in-house technician help line for assistance. Kodak claimed such sales were unintentional.

Over time, Kodak increasingly policed its no-parts sales policy. It began tracking parts sales to ensure that equipment owners were not purchasing more than they would reasonably need for the equipment they owned. Kodak required prospective parts purchasers to provide proof of ownership of the equipment model for which they were ordering parts. It additionally required, for copier parts, certification that the customer had a Kodak-trained employee to effect the repairs. Kodak also required customers to agree to not resell parts.

PROCEDURAL HISTORY

The Initial Case and the Motion for Summary Judgment

In April 1987, ITS and several other ISOs filed suit against Kodak in the District Court, Northern District of California. The suit alleged that Kodak used its monopoly over parts to monopolize the service markets for its copiers and micrographics equipment; that Kodak had conspired with its

*The record is not clear on whether the policy was implemented in 1985 or 1986. It was not much enforced, if at all, prior to mid-1986.*
outside parts suppliers to preclude ISO access to parts, thereby monopolizing the service market; and that Kodak had tied its service labor to parts, thereby harming competition in the service market. These actions were alleged to violate Sections 1 and 2 of the Sherman Act.

Very early in the fact-finding process, Kodak moved for "summary judgment." Kodak argued that there was no allegation that it had market power in the equipment markets. Kodak claimed that equipment customers had many alternatives available to them and made purchase decisions based on the total cost of ownership, and thus any attempt by Kodak to extract higher profits from maintenance customers would result in equipment customers' taking their business elsewhere. Thus, because it could not have service market power, Kodak argued that as a matter of law it could not be found guilty of tying or monopolizing service markets. The judge agreed and granted summary judgment.

The ISOs appealed to the Ninth Circuit Court of Appeals, claiming that Kodak's arguments were purely theoretical. Various market imperfections could break the link between higher aftermarket prices and reduced foremarket sales, preventing foremarket competition from sufficiently disciplining aftermarket market power. The Ninth Circuit granted the appeal, overturned the summary judgment order, and remanded the case for trial. Kodak appealed to the Supreme Court, which ruled in 1992. The Court agreed with the Ninth Circuit and remanded the case for trial in district court.

It is important to recognize the Supreme Court did not find that Kodak had illegally monopolized or tied. Nor did it find that parts or service were necessarily relevant markets. The Supreme Court found only that Kodak had not carried its burden of proof for summary judgment. Kodak had failed to convince the Court that economic theory proved that it was impossible for Kodak to be guilty. Rather, the Supreme Court concluded that the ISOs' economic theories were plausible and that Kodak's guilt or innocence hinged on the interpretation of the facts in the case: A trial would be required.

The Trial

On June 19, 1995, the jury trial of ITS v. Kodak began in district court. The trial involved sixty-three witnesses giving twenty-seven days of testimony.

"Summary judgment" is granted when the court is persuaded that, even if all disputed facts are resolved in favor of the nonmoving party, the nonmoving party cannot, as a matter of law, win the case. Essentially, it is a way to avoid wasting court resources.

"Total cost of ownership" is the present discounted (expected) value of all costs associated with owning and using the machine, including future maintenance costs. This is also called "life cycle" analysis.

1988WL 156332 (N.D. Cal.).
903 F.2d 612 (9th Cir. 1990).
At the close of evidence, the plaintiffs dropped the tying claims. Therefore, the only claims presented to the jury were that Kodak had monopolized the service markets for its high-volume copier and micrographics equipment. The alleged monopolizing acts were the restrictive parts policies, which leveraged Kodak’s parts monopolies into service monopolies. The jury deliberated for thirteen days and returned a unanimous guilty verdict. Kodak was held liable for $24 million in damages, trebled to $72 million. On February 15, 1996, the district court issued a ten-year injunction requiring Kodak to sell parts to ISOs at nondiscriminatory prices.12

Kodak appealed to the Ninth Circuit Court of Appeals with three interesting economic points: Can Kodak be required to sell patented parts and copyrighted service software and manuals? Can “all Kodak parts” be a relevant market, despite the lack of substitutability between two different parts? Can a firm be convicted of monopolizing its aftermarkets without first being found to have obtained supracompetitive systems profits or prices?7

On August 26, 1997, the Ninth Circuit ruled. The court rejected these three Kodak arguments and upheld the plaintiffs’ verdict on all liability issues.

ECONOMICS OF AFTERMARKET MARKET POWER

The central economic feature of Kodak was the dispute over theories of market power in aftermarkets. Kodak argued before the Supreme Court that primary market competition, as a matter of economic theory, precludes anticompetitive aftermarket actions. After the Supreme Court rejected this argument, Kodak argued at trial that the presumption should be that aftermarket power was unlikely and that Kodak’s circumstances were consistent with this presumption. We refer to this as the “systems” theory. The plaintiffs argued that there are several theories showing commonplace circumstances under which a durable goods manufacturer could monopolize its aftermarkets. Therefore, they argued, there should be no presumption against aftermarket power. Further, the plaintiffs presented facts to demonstrate that Kodak both had and exercised market power. To present the arguments put forth on each side, we will use a very simple model of firm profits from a foremarket and an aftermarket.

Whereas a firm competing in a single good has only a single price lever, a firm participating in fore- and aftermarkets considers both prices


13The court required a new trial to recalculate a portion of the damages because the plaintiffs’ accounting expert and the fact witnesses did not sufficiently link some of the damages to the antitrust violation. Image Technical Services, Inc., et al. v. Eastman Kodak Co., 125 F.3d 1195 (9th Cir. 1997).
The firm chooses these to maximize profits \( \Pi \), which, on the assumption of constant costs \( (c_f, c_a) \), are:

\[
\Pi = \pi^f + \pi^a = (p^f - c^f)q^f(p^f, p^a) + (p^a - c^a)q^a(p^f, p^a).
\]

We assume that the quantities demanded for each product depend on both prices, for example, \( q = q(p^f, p^a) \). That is, if the price of service rises, we generally expect the demand for primary market equipment to fall. This linkage between goods across markets is the essential feature of aftermarket economics. Most of the disputes can be summarized as arguments over how strong that link is under various factual circumstances.

We start by assuming that foremarkets and aftermarkets are perfectly competitive: \( p^f = c^f \) and \( p^a = c^a \). We are interested in whether a firm can increase its profits by acquiring market power in its aftermarket while its foremarket remains competitive. Thus, we are asking if a firm can profit overall by raising \( p^a \) above \( c^a \). That is, the basic disagreement is whether, starting from competitive pricing, the sign of \( \frac{\partial \Pi}{\partial p^a} + \frac{\partial \Pi}{\partial p^f} \) is positive or negative.

Economists on both sides agree that under most circumstances aftermarket profits can be increased, \( \frac{\partial \Pi}{\partial p^a} < 0 \) but that there will be a simultaneous decline in foremarket profits, \( \frac{\partial \Pi}{\partial p^f} < 0 \), because increasing \( p^a \) increases the overall cost of owning equipment and forward-looking buyers will reduce their equipment purchases. The debate, then, is about whether \( \frac{\partial \Pi}{\partial p^a} \) is sufficiently negative to offset monopoly aftermarket profits and thus make aftermarket monopolization unprofitable overall.

Economists writing since the Supreme Court opinion, including at least one who testified for Kodak at trial, have largely agreed that there are circumstances under which aftermarket monopolization can be profitable overall; that is, \( \frac{\partial \Pi}{\partial p^a} > -\frac{\partial \Pi}{\partial p^f} \). We shall briefly describe some of the theories in order to frame the facts and economic arguments presented at trial.

Footnote 14: Whether firms compete in prices ("Bertrand") or quantities ("Cournot") is not important here; we discuss "price" competition for convenience. However, focus on just price alone (or just quantity) is an important simplification. Consumers consider factors beyond just price when products are differentiated. Such differentiation is particularly true with "service" type goods: Two service programs might differ in technical quality, response time, flexibility of contract terms, provision of manufacturer-independent advice, and so forth. Kodak surveys showed that typical customers rated several factors as more important than price when choosing a service provider (see Ex. 264).

Footnote 15: Typically, aftermarket sales associated with a given foremarket sale will occur over a period of time. Therefore, what we refer to as aftermarket prices (for consumers) and profits (for firms) are actually discounted streams of future expenditure and profits, respectively.

Footnote 16: See, e.g., Borenstein, MacKie-Mason, and Netz (1995, 2000), Shapiro (1995), Carlton (2001), Chen and Ross (1993), and Waldman (2007). There is yet another issue, which is that under some circumstances even if it is profitable overall to monopolize an aftermarket, social welfare may be harmed only a little, or even improved, and thus legal antitrust intervention might be inappropriate. See footnote 26 below.
Preliminaries

Under what conditions might we expect foremarket competition to protect aftermarket consumers? That is, when might the total effect, \(3jt/dp\), be zero or negative, so that aftermarket monopolization is not attractive?

First, for the question to be interesting there must be at least some aftermarket power, which we define as \(3jr/3p > 0\). That is, the firm must have the ability to raise the aftermarket price above the competitive level and earn additional aftermarket profits, ignoring the effect on foremarket profits. This requirement implies two necessary conditions: There must be some protectable aspect to the aftermarkets and some form of switching costs related to the primary market good.

Protectable aftermarkets mean that there are limited substitutes for, and limited entry into, the equipment manufacturer’s aftermarket. If aftermarket substitutes are widely available at competitive prices, say, from independent parts manufacturers or service providers, increases in \(p\) will merely result in the firm’s equipment customers making their aftermarket purchases from another supplier. That would imply \(3n/3p \leq 0\).

Similarly, if equipment owners can switch at no cost from their existing equipment to a competitively priced alternative, a service price increase would induce consumers to sell their equipment and buy the alternative brand of equipment and service. It must be costly for current equipment owners to switch to another brand of equipment if the manufacturer is to have aftermarket power. There are two basic types of switching costs: inefficient used equipment markets and complementary sunk investments that are specific to the given brand of primary good. An inefficient used equipment market means that the seller of a used primary market good cannot expect to recover the full economic value of the good when she sells it. This could be due to a lemons problem or to other causes. We will refer to these as “financial” switching costs. Financial switching costs are coterminous with the economic life of the specific piece of equipment: Once a machine has zero economic value, there is no financial cost of switching to a new brand.

Switching costs from complementary investments arise when a firm needs to make investments in addition to the equipment and aftermarket good in order to utilize the equipment and these investments are of little value with any other brand of equipment. The classic example of complementary investment is custom applications software written for a specific operating system—switching to a new operating system requires rewriting.

\[ ^{17} \text{This implies that the primary market is monopolistically competitive or a differentiated product oligopoly. It is difficult to conceive of protectable aftermarkets in combination with truly homogeneous primary market goods: If there are no differences among primary market goods, how can a given aftermarket good work only with one “brand” primary market goods?} \]

\[ ^{18} \text{See Akerlof (1970). However, even a perfectly functioning used-equipment market might not eliminate financial lock-in resulting from installed base opportunism. See the text below.} \]
the applications. Other switching costs include training and familiarization, converting data and archival file formats, custom configuration of peripherals, developing new relationships with expert, sales, and service personnel, and so forth. We refer to these as "technological" switching costs. Technological switching costs can extend beyond the economic life of an individual piece of equipment."

The level of total switching costs puts an upper bound on the amount of surplus a firm can extract by exercising aftermarket power. Any attempt to extract the current owner's surplus in excess of the cost of switching results in the equipment owner's switching brands. However, there are situations where switching costs are large and this constraint might not be significant. When switching costs are significant, we say consumers experience lock-in."

If aftermarkets are not protectable and switching costs are negligible, then we would generally expect no local aftermarket power. $\frac{\partial \delta}{\partial p^a} < 0$, and the systems theory would prevail.

**Necessary Conditions for the Systems Theory**

When a durables manufacturer has local aftermarket power ($\frac{\partial \delta}{\partial p^a} > 0$), systems theory proponents argue that linkage to a competitive foremarket will protect locked-in aftermarket consumers from monopoly exploitation. That is, $\frac{\partial \delta}{\partial p^a}$ is sufficiently large and negative so that a foremarket profit loss offsets the aftermarket profit gain. The argument is either that switching costs are low, so that current equipment owners can switch at low enough cost to constrain sufficiently the manufacturer's aftermarket power, or that new customers (and repeat customers upgrading or replacing their equipment) will see the high aftermarket prices and demand lower equipment prices to compensate, or will purchase elsewhere. What conditions are necessary to establish a sufficiently strong linkage between fore- and aftermarkets? Three have received the most attention: sufficient, low-cost information; effective simultaneity of fore- and aftermarket purchases; and competitive foremarkets for equipment. We explain each in turn.

**Sufficient, Low-Cost Information**

The systems theory assumes that consumers are aware of aftermarket prices, make reasonable assumptions about their own future demand for the

---

19 This can lead to the confusing situation where future equipment purchases should properly be analyzed as aftermarket purchases: The choice of which primary market good to buy in the future is contingent on which is purchased today. Thus, technological switching costs increase competition in new product markets but reduce competition in mature markets. With new markets, firms compete vigorously to get a large base of locked-in customers; in mature markets, most customers are already significantly locked in to a brand. See Klemperer (1987).

20 For more on switching costs and lock-in, see Porter (1985), Williamson (1985), and Farrell (1985).
aftermarket goods corresponding to each primary market good, and use this information when making price comparisons. Thus, consumers are aware when a given primary market supplier charges supracompetitive prices in the aftermarket and take their custom elsewhere.

**Effective Simultaneity**

Aftermarket goods and services are typically purchased later than foremarket equipment. Yet, the systems theory requires that an increase in aftermarket prices be accompanied by an offsetting decrease in foremarket profits. This linkage requires what we call "effective simultaneity"; that is, the markets must operate as if the fore- and aftermarket purchasing decisions were being made simultaneously. Systems theory proponents suggest several ways in which effective simultaneity might be obtained. First, product lifetimes might be sufficiently short that current equipment owners will soon be purchasing new equipment and will expect that, if charged high current aftermarket prices, they will be charged high prices in the future. (Note that this argument assumes low or nonpersistent technological switching costs, since these can lock in repeat purchasers beyond the life of their current equipment.) Second, the ratio of new potential buyers to existing owners may be sufficiently large that high current aftermarket prices will dissuade sufficient new buyers, because they expect to be charged high aftermarket prices. (This argument requires that the firm not be able to price discriminate openly between new and old buyers.) Both approaches require that a firm's pricing reputation is important because consumers consider the firm's past pricing.

Another approach to obtaining effective simultaneity is through warranties and long-term aftermarket contracts purchased at the time of equipment purchase. Similarly, the availability of rental or lease agreements that include the aftermarket good or service might restrain a firm's aftermarket power by reducing lock-in.²¹ These arguments received little attention in Kodak (and similar cases) because aftermarket contracts extending for the full life of the equipment were never offered by Kodak.

**Competitive Foremarkets**

The third premise of the systems theory that durable manufacturers cannot have aftermarket power is that the equipment foremarket be competitive. The linkage argument requires that an attempt to charge high aftermarket prices will be foiled by strong competition in the foremarket.

**Theories of Market Power in Aftermarkets**

Kodak proposed to the Supreme Court in 1992 that as a matter of law, on the basis of the systems theory, it should be held that it was not possible

²¹However, these protect against only financial, not technological lock-in.
for Kodak to have aftermarket power, and thus the case should be dismissed without a full factual inquiry. Various economists—some involved in the case, some not—responded by showing that under plausible circumstances any of the necessary conditions of the systems theory might not hold and that the conclusion about lack of market power in the aftermarkets would then change. We shall now briefly describe the main theoretical challenges to the systems theory.

**Installed Base Opportunism (IBO)**

One theory disputes the assumption of effective simultaneity. For at least some customers, there is not simultaneity: those who already own equipment and who are not about to replace it. Even if the other assumptions hold, the manufacturer could practice what is known as installed base opportunism: After customers are locked in, surprise them by changing policies in a way that raises aftermarket prices above the competitive level. For example, Kodak changed its policy of making parts available to independent service providers. Competition in the foremarket may force the firm to discount new equipment sales to offset the aftermarket price increase, thereby earning zero economic profits on new sales, but due to competition the firm was already earning zero profits on all customers. Now the firm gets excess profits equal to the total switching costs faced by its installed base. IBO will be especially attractive for mature and declining product lines, in which most revenues are from locked-in, rather than new, customers.

**Costly Information**

Consumers will rationally forgo complete life-cycle cost analysis if sufficient low-cost information is not available. Costly information also mitigates the effectiveness of reputation and thus undercuts the effective simultaneity assumption. Since maintained durables often have lives of seven, fifteen, or even more years and may have parts lists thousands of items long, obtaining sufficient information about future costs, and future user needs, can be quite difficult. In addition, both durables and their aftermarket goods and services may be highly differentiated. This means a complete analysis requires extensive information gathering about each of the various models and brands considered, as well as about many features other than price. Indeed, many aftermarket products—for example, maintenance or software—have significant "experience" components: Potential buyers have great difficulty knowing their value ex ante.

---

22 Salop (1992) applied IBO theory to Kodak. The theory had been applied previously in other contexts. See, e.g., Williamson (1985).

23 The plaintiffs presented this theory to the Supreme Court; Shapiro (1995) has discussed it.

24 For some background on experience goods, see Carlton and Perloff (1994, pp. 601-602) and references and Tirole (1988, pp. 294-295) and references.
Imperfect Foremarket Competition

Another critique of the broad applicability of the systems theory relaxes the assumption of perfectly competitive equipment markets. Markets for expensive durable equipment are often quite concentrated; and even if they are not, there is often a substantial degree of product differentiation. Thus, we see the conditions for monopolistic competition, with each firm facing a downward-sloping residual demand curve."

Klemperer (1987) has shown that in an oligopolistic market with switching costs the collusive output level can occur in a noncooperative equilibrium. When foremarkets are not competitive, it will be generally true that a manufacturer can increase its profits further by monopolizing the aftermarket (see, e.g., Borenstein, MacKie-Mason, and Netz 2000). Several authors have pointed out that one way in which aftermarkets can be used to increase foremarket profits is through price discrimination (for example, Chen and Ross 1993). Suppose that customers are heterogeneous in the utility they derive from the good and its aftermarket. This heterogeneity is not observable in the primary market. The heterogeneity is, however, related in some manner to the consumer’s demand for the aftermarket good. This allows the equipment manufacturer to use aftermarket purchases as a metering device if he has an aftermarket monopoly, and thereby to extract additional profits."

Imperfect Commitment

Borenstein, MacKie-Mason, and Netz (1995) present a model that is based on precisely the assumptions of the systems theory and show that even then there will generally be supracompetitive aftermarket pricing and harm to consumer (and overall) welfare. The main point is that when firms cannot sign complete contracts for aftermarket products that cover the full life of the durable good, and all contingencies, then effective simultaneity will generally be an approximation, and there will be at least some room to profit from the slippage.

The intuition is not hard: As soon as the firm has some locked-in customers, it can earn at least some monopoly profits by charging a supracompetitive aftermarket price. Due to foremarket competition it may lose some

"In high-volume copiers, Kodak and Xerox produced about 90 percent or more of all units sold during the relevant years. Not only was the market concentrated, but also the products were substantially differentiated.

"Chen and Ross (1993) have modeled the use of aftermarkets for price discrimination. They note that price discrimination itself has ambiguous welfare effects: The firm unambiguously increases its profits, but combined producer plus consumer welfare can either increase or decrease. Indeed, Hausman and MacKie-Mason (1988) have shown that price discrimination can even lead to a Pareto improvement if new markets are opened or economies of scale are significant. However, these concerns do not undermine the basic point: When a firm has foremarket power and can use the aftermarket to meter, it will earn higher profits due to aftermarket monopolization, contrary to the systems theory. It is still necessary to show antitrust injury from the profitable monopolization effort before winning a case."
foremarket profits and need to lower equipment prices, but the firm was earning only competitive (zero) profits before, and the cost of a small foremarket price change at the margin is zero, or small compared with the aftermarket profit gain. That is, given switching costs and lock-in, a rational firm will always want to extract at least some of the aftermarket monopoly profit, even if foremarket competition prevents it from extracting all of the profit. The result does not require the surprise element of the IBO theory.

The consumer harm could be much larger if, for example, aftermarket monopolization reduces desirable product differentiation (such as variety in service terms or quality), or eliminates innovation that might have occurred in a competitive aftermarket (Borenstein, MacKie-Mason, and Netz 1995). Since this theory shows that the systems theory never holds completely and that the magnitude of harm depends on the facts, it directly challenges the view that aftermarket monopolization should be strictly or even presumptively legal as a matter of law."

THE ARGUMENTS PRESENTED IN THE SUPREME COURT AND IN TRIAL

Supreme Court (Summary Judgment 1992)

It bears repeating that Kodak first came to the Supreme Court as the result of Kodak's motion for summary judgment. This has two implications for the outcome: All questions about disputed facts were resolved for the nonmoving party (the ISOs), and the moving party had to demonstrate that it could not be found guilty under any reasonable interpretation of the facts of the case. As the Supreme Court noted in its decision, the motion was heard early in the discovery phase of the case: The Court had a slim factual record upon which to base its decision.

Kodak argued that it faced substantial equipment competition in both copiers and micrographic equipment. Kodak claimed that it competed with Xerox, Canon, and others in copiers. It listed Canon, Anacomp, Bell & Howell, 3M, and others among its micrographics equipment competitors. Kodak argued that, given this equipment competition, any attempt by it to abuse its aftermarket customers would have a ruinous impact on its equipment sales: It could not profitably exploit whatever market power it had in its

"European Union law essentially presumes aftermarket monopolization is harmful—firms are required to sell parts. See Reed (1992); Waldman argues in a series of papers (summarized in Waldman 2007) that the factual conditions for non-IBO information problems or limited commitment are not "likely," which (of course) is an empirical question. The facts in Kodak, developed during a full trial, demonstrated that consumers had incomplete information about basic aspects of aftermarkets at the time of foremarket purchases. Waldman also offers an alternative theory in which aftermarket monopolization by a firm with no foremarket power can increase consumer welfare by internalizing and thus making efficient the choice between replacing and maintaining durable equipment.
aftermarkets. Therefore, parts and service could not be distinct relevant mar­
kets for antitrust, and Kodak could not be guilty of tying or monopolization.

Under the case law, even if Kodak tied or monopolized, it might not
be guilty if it had legitimate business justifications. Kodak asserted three
procompetitive business justifications: desire to provide quality mainte­
nance, desire to control its inventory costs, and desire to prevent the ISOs
from "free-riding" on Kodak’s investments in equipment, parts, and service.

The ISOs countered that customers were not perfectly informed and
that they faced high switching costs. Therefore, even if equipment markets
were competitive, aftermarket customers might not be protected from
abuse. The ISOs also argued that there was evidence that Kodak had
engaged in IBO. Thus, they relied on two of the theories that undo the sys­
tems theory.

The Court rejected Kodak’s argument in a detailed opinion. The Court
noted that Kodak’s theory required factual assumptions about the real
world;"" Kodak’s theory was not compelling on purely theoretical grounds;"" and Kodak’s theory, rather than being supported by evidence in the record,
was contradicted by the record."" Furthermore, the ISOs presented plausible
explanations for why Kodak’s theory didn’t explain the evidence."

**Trial**

The case at trial was somewhat different than the case argued before the
Supreme Court. This difference stemmed in large part from the plaintiffs’
discovery that Kodak had substantial market shares in what were very con­
centrated foremarkets.

Although this case involved copiers and various micrographics equip­
ment, we will concentrate on the analysis of copiers. At the end, we shall
briefly describe some differences in the micrographics part of the trial.

**Plaintiffs’ Arguments**

The plaintiffs made three main arguments at trial: first, that Kodak had
monopoly power over repair parts and leveraged this power to maintain and
extend a service monopoly; second, to rebut Kodak’s main defense, that

""Kodak’s proposed rule [that “equipment competition precludes any finding of monopoly power
in derivative aftermarkets” (504 U.S. 451 at 466, citing Kodak’s Brief at 53)] rests on a factual
assumption about the cross-elasticity of demand in the equipment and aftermarkets . . . ."" Kodak,
504 U.S. 451 at 469. That is, the Court recognized that the strength of the linkage depended on the
factual circumstances of the case.

""Thus, contrary to Kodak’s assertion, there is no immutable physical law—no ‘basic economic
reality’—insisting that competition in the equipment market cannot coexist with market power in
showed later that even Kodak’s idealized conditions, stated in Justice Scalia’s dissent, imply at least
some market power in aftermarkets.

*Kodak, 504 U.S. 451 at 472.

*Kodak, 504 U.S. 451 at 473.

Page 571
Kodak in fact had significant market power in the foremarkets; third, also in rebuttal to Kodak, that several factors sufficiently broke any linkage between fore- and aftermarkets.

**Market Definition:** The plaintiffs argued that there was a relevant market for the repair parts needed for Kodak copiers. The evidence indicated that many Kodak copier parts had no substitutes at all: They were unique to Kodak copiers. Furthermore, there were substantial costs involved in switching to another high-volume copier brand, indicating that customers would have an inelastic response to a significant increase in the price of Kodak copier parts. Last, there were barriers to entry into production of Kodak parts. Evidently a single firm controlling access to the parts could profitably raise prices substantially.

The “all parts” nature of the parts market definition was a departure from a strict interpretation of the DOJ/FTC Guidelines methodology. It resulted in clustering complements within the same relevant market. For example, fuser rollers and image loops were both in the relevant market. The plaintiffs argued that “all parts” were a relevant market based on the “commercial realities” of the case: If parts demanders could not get, for example, fuser rollers, they would have no demand for image loops. Kodak’s parts manager testified that one would need “an assured supply of parts” to be in the service business. Most customers would deal only with service providers who had an assured source of supply for all parts. Additionally, Kodak’s policy applied to all parts, not just specific parts.

The switching costs that were relevant for parts were equally relevant for service labor, as were the theories undercutting the systems theory. Thus, a single firm controlling all service could profitably raise price substantially, and service was a relevant market.

Last, the plaintiffs argued that there was a relevant market for high-speed, high-volume (HV) photocopiers. Customer testimony, internal Kodak documents, and industry sources all strongly suggested that there was a distinct HV copier market and that it was a duopoly composed of Kodak and Xerox. This led the plaintiffs to define the HV copier market using copier speed, volume capability, and durability measures.

**Market Power:** The plaintiffs contended that Kodak had market power in the parts market. Kodak had patents on various critical parts and refused to provide the specifications on others. There were entry barriers to parts production apart from the patents—in particular, significant minimum

---

32 Tr. at 5585. Kodak used 9942 different parts in one year to service its installed base of copiers. It used 6430 different parts for servicing its micrographics equipment. Tr. at 5558-5559.

33 The plaintiffs used two measures of copier speed (reproducing a single-sided, single-page document and a two-sided, multipage document), rated monthly copy volume, and weight (to reflect durability). Though some Japanese copiers with high single-page speeds were offered by the early 1990s, these units still fell far short of the Kodak and Xerox offerings in the other characteristics. See Tr. Ex. 724, 725, 726, 744, 752, 753.
efficient scale.” Because a potential service competitor required access to all parts to compete, these restraints were sufficient to conclude that Kodak had monopoly power in the parts market.

Kodak also allegedly had market power in the service market. Kodak’s market share was 98 percent.” There was evidence that Kodak restricted service contract options” and engaged in significant price discrimination across service customers, both of which are indicators of possible market power.”

The plaintiffs also argued that the equipment market did not discipline the exercise of market power in the aftermarkets. One significant piece of evidence was that Kodak did not engage in systems pricing. Where systems pricing requires that customers paying above-average service prices pay below-average prices for their equipment and vice versa, there was essentially no statistically significant correlation between the prices customers paid for equipment and the prices they paid for maintenance. The data showed that customers paid widely different systems prices on a total-cost-of-ownership (TCO) basis.”

Plaintiffs gave several reasons for this unlinking:

• There were significant switching costs across copier brands.” In addition, shortly after introducing its parts policy Kodak increased switching costs within the Kodak service market by quadrupling the inspection and restoration fees that it charged for Kodak to begin servicing a copier bought from or previously serviced by someone other than Kodak.”

• Several equipment customers testified that they did not engage in TCO. A Kodak saleswoman indicated that her “TCO” proposals for customers included at most an initial few years of service while the equipment generally lasted much longer.” Kodak’s expert testified that the inclusion of later renewal agreements was necessary for TCO calculations to be correct.

---

34 A copier parts manufacturer testified that, for most parts, he would need to be supplying parts for 2000 or more machines, roughly 5 percent of Kodak’s installed base, to generate a sufficient return on design and tooling investments (ignoring technical and patent difficulties). Tr. at 818-820, 825-827.

35 Ex. 730. This left the ISOs with one-fourth to one-twentieth the share held by ISOs in some other high-tech industries. Kodak’s expert said Kodak’s share was 90 percent. Tr. at 4573:11-16.

36 Ex. 215. For example, ITS offered CSC an on-site technician, but Kodak did not.

37 Exs. 238, 239.

38 SeeEx. 743.

39 Ex. 735. See also Plaintiff’s Expert Report on Liability Issues for a more detailed analysis.

40 Ex. 742. This action made it riskier to try ISO service: if you were not satisfied with the ISO, it would be very expensive to return to Kodak service.

41 Tr. at 4511-4517.

42 Tr. at 4587-4589.
• Kodak had market power in the equipment market. Kodak and Xerox were essentially the only suppliers in the high-volume copier equipment market. There was also extreme price discrimination across equipment purchasers.

The plaintiffs showed that when ISOs could obtain access to parts they could compete with Kodak for service contracts with lower prices and comparable quality. Kodak’s own documents recorded that the parts policy helped Kodak exclude competition (see the quotes at the opening of this chapter). Therefore, the plaintiffs argued, Kodak’s parts policy was used to maintain and extend its service monopoly.

**IBO:** Installed base opportunism was a prominent, though not the only, theory presented to the Supreme Court. It was not the central theory at trial for copiers, but the plaintiffs did present IBO evidence. Such evidence was more prominent for micrographics. Several customers testified that after a few years on Kodak maintenance, the price increased significantly. Additionally, a Kodak saleswoman testified that renewal maintenance agreement terms were more limited and less attractive than for agreements purchased at the time of equipment sales. Thus, although there was not a specific date on which Kodak raised prices to all current owners of Kodak equipment, there was evidence that IBO was applied customer-by-customer. Further, earlier standard industry practice had been to permit independent service. Both IBM and Xerox had sold parts for their copiers. Kodak had sold parts for its micrographic equipment. Therefore, when Kodak implemented a policy to refuse parts sales, and thus stunted the development of ISOs as the market matured, customers were surprised to find that ISOs could not get parts and that prices were higher than they would have been, even if they did not rise in nominal terms.

**Kodak’s Arguments**

Kodak’s expert testimony closely paralleled its Supreme Court case. This meant that it followed a significantly different framework than the plaintiffs’ presentation. Kodak asserted the presumption that equipment competition precluded profitable abuse of aftermarket power and that therefore “systems” were the relevant market.

**Market Definition:** Consequently, Kodak’s market definition effort focused strictly on the copier equipment market. It claimed that the market

---

43 Ex. 748. The Herfindahl index for high-volume copiers was over 3100.
44 Some customers paid 1.7 to 2.3 times as much for a given equipment configuration as other customers purchasing the same quantity.
45 Tr. at 4511-4517.
46 ISOs, when they could get parts, offered service at about 15-30 percent less than Kodak, and when Kodak competed head-to-head it cut prices drastically. See, e.g., Ex. 729.
was for copiers with a multicopy speed greater than 60 cpm. By ignoring throughput and durability, Kodak argued there was significant competition in the 1990s. Kodak also argued that the plaintiffs' price discrimination evidence did not indicate market power in a concentrated market, but rather hard customer bargaining in a vendor-competitive environment."

**Market Power in the Aftermarkets:** Kodak's experts then addressed whether this competitive equipment market would adequately protect Kodak's aftermarket customers. They presented hypothetical examples to suggest that Kodak would lose profits from lost future systems sales if it overcharged aftermarket customers." They asserted that it was reasonable to expect that Kodak would lose future sales if it exploited aftermarket power because a significant part of both its micrographic and copier sales was made to existing customers," but they offered no direct evidence that repeat purchasers were responsive to aftermarket practices.

Kodak's copier expert suggested that switching costs were not significant. He did not analyze them in detail because he had concluded that it would not be profitable for Kodak to exploit lock-in even if it did exist.

Kodak argued that information costs were also not significant. It presented evidence that several publications were available that gave guidance on how to conduct TCO as well as giving independent, if anecdotal, pricing information. Additionally, information cost effects are mitigated by the high proportion of repeat purchasers: Because such customers would be able to spread any costs of performing TCO over their several purchases, they would have lower information costs per equipment unit. Kodak's experts also noted that Kodak had spent considerable money developing an automated quote system that included pricing on initial service contracts. They argued that Kodak wouldn't have done this if either it was counting on customer ignorance or customers hadn't regularly asked for such TCO information.""

Kodak also pointed out that equipment manufacturers typically have large aftermarket shares. Kodak "invented" the Kodak aftermarket,

---

"Tr. at 4634, 4669-1670.

"Kodak's economic expert on copiers believed that the systems theory was dispositive: He testified that merely knowing that Kodak had to compete with Xerox would have been sufficient for him to conclude that Kodak could not harm its aftermarket customers. Tr. 4515-23=1516:3.

"Ex. 3697 and Tr. at 4663-1608. However, the exhibit was based on the assumption that Kodak would lose all future systems sales to this customer. This is the exact assumption that the Supreme Court refused to accept as dispositive, demanding instead that the facts be examined. Kodak presented a similar exhibit in its micrographics presentation, using the same underlying assumption. Ex. 3768 and Tr. at 6018-6014, and 6323-6325.

"Seventy-five percent of Kodak's copier sales between 1986 and 1994 were to past or current owners of Kodak copiers. Ex. 3695. In micrographics, somewhere between 68 percent and 90 percent of purchases from 1990 to 1994 were to past or current owners of Kodak equipment. Ex. 3733 and Tr. at 5999-6000.

"This conclusion was based on Ex. 3697. See fn. 48 and related text.

"Tr. at 4592-4593."
inherently having a 100 percent share at the inception of the aftermarket. Thus, Kodak’s “monopoly” share in the market was not the result of Kodak’s parts policy but was instead a natural effect of Kodak’s having created the market."

**Business Justifications:** Kodak claimed several pro-competitive business justifications for its parts policy. Kodak argued it had chosen to compete in the equipment market by adopting a strategy of providing high-quality maintenance. This and the avoidance of finger-pointing were reasons to prevent ISOs from servicing Kodak equipment. Kodak claimed that it also needed to be the sole service provider so that it could control its parts inventory costs. Kodak argued that the ISOs were free-riding on its investment in developing maintenance methods, tools, and parts. Last, Kodak argued that its patents gave it the explicit right to refuse to sell patented parts.

**Jury Findings**

The jury returned a unanimous verdict against Kodak. Based on the written instructions from the judge to the jury (and on the written opinion of the Ninth Circuit, which reviewed the case), we can infer that the jury concluded:

1. Equipment, parts, and service were distinct relevant markets.
2. Kodak had a parts market monopoly.
3. Kodak levered its parts monopoly to monopolize service.

**ANTITRUST AND ECONOMICS AFTER KODAK**

*Kodak* raises interesting issues about the role of economics in antitrust analysis that transcend the economics of aftermarkets. We shall briefly describe two here; space is too limited for a thorough treatment. First, should a plausible but relatively untested economic theory be sufficient, as a matter of law, to prevent a case from getting to a factual determination by a jury? Second, having gotten to the jury, where does the plaintiff’s burden

---

*Kodak’s expert presented evidence on other industries, including IBM copiers, in which ISOs had only a 5 percent market share. The IBM copier example is particularly apropos because IBM was required to sell parts for its equipment to anyone who asked. Tr. at 4578, Exs. 3693, 3692.

*Finger-pointing was described as the customer’s inability to distinguish between poor service and poor equipment. Thus, Kodak claimed, it had the right to control service because poor service could harm Kodak’s equipment reputation.

*Juries, unlike judges, do not issue written opinions to support their decisions, but the jury was instructed by the judge that it had to make specific findings on several questions before it could conclude that Kodak was guilty.
of proving the relevant market with partial equilibrium evidence in a general equilibrium world end?

Legal intervention in markets, even when intended to correct market failures, is itself costly and subject to error. The summary judgment procedure exists in part to balance the costs of intervention with the benefits of antitrust enforcement. A party can move for summary judgment on the basis that it will prevail as a matter of fact, even if the facts are all interpreted in the other party’s favor. Kodak made a somewhat novel motion for summary judgment: It claimed that on the basis of an economic theory, a showing that Kodak had market power would be impossible even if the plaintiffs had the facts right. If the plaintiffs could not establish market power, Kodak would prevail. The question raised is whether courts should rely on a novel, untested economic theory to conclude that despite the facts, the plaintiffs would not be able to demonstrate market power.

When dealing with questions such as market power, courts necessarily rely on economic theories. However, it is our view that there should be a strong presumption against granting summary judgment on the basis of a theory that has not been well tested in previous cases, or for that matter in the economics literature. Kodak’s theory had a superficial plausibility and was endorsed by Justice Scalia in his dissent to the majority Supreme Court position. However, as shown above, economic research since Kodak has established that the plaintiff’s alternative theory is sound, and additional theories have been published and accepted under which market power in an aftermarket could exist.” When an opposing party offers a plausible alternative theory, judges are called on to evaluate the competing theories for correctness and appropriateness to the alleged factual setting, and since this is not their area of expertise, the risk of judicial error is high.

The second question, we think, is fundamental to the practice of antitrust economics: When it is impossible to do a complete, general equilibrium analysis of market interactions, how far does the burden of proof on the plaintiffs extend? Kodak argued in trial and in its appeal to the Ninth Circuit that since higher service profits might be offset by lower equipment profits under systems competition, the plaintiffs were obliged to show that combined profits were above competitive. Only then, Kodak argued, would the ability profitably to monopolize a service market be established; that is, only if Kodak did earn combined monopoly profits would plaintiffs have shown that it had market power. This is a difficult proof issue that we believe requires a policy judgment. Should it be the plaintiff’s burden to prove that systems competition did not protect consumers from antitrust harm, or should it be the defendant’s responsibility to prove that systems competition did protect consumer welfare?

*Indeed, one of Kodak’s expert economists at trial has published an article acknowledging the theoretical correctness of four of the alternative theories (Shapiro 1995).*
In an ideal world, we would assess all interactions between a hypothetical monopolization and other markets. If, taking into account all interactions, it would not be possible for a defendant to harm overall social welfare, then we would not find it guilty of an antitrust violation. A complete, general equilibrium analysis will almost never be possible, yet clearly it is not consistent with antitrust policy to bar all plaintiffs from court because they are unable to perform such an extensive analysis. In general, all market definition analyses are partial equilibrium in nature.

Consider an example. Suppose Kodak asserted that by charging supracompetitive service prices, enough disgruntled customers would stop buying Kodak film products that Kodak overall would not be able to earn supracompetitive profits. (Most micrographic equipment in this case requires film, and Kodak, of course, also sells other film.) Should plaintiffs also have the burden to prove that film profit losses do not outweigh service profit gains? What if Kodak proposed an even more remote linkage? Where should the line be drawn?

The question raised about burden of proof and the proper bounds for practicing market definition apply generally. Plaintiffs practice and courts permit partial equilibrium analysis. However, this provides only general guidance. How far the plaintiff has to go is a policy question for Congress or the courts.

WHITHER KODAK AND INDEPENDENT EQUIPMENT SERVICE?

By the time of the circuit court opinion in 1997, Kodak’s photocopier business was clearly in decline. In late 1996 Kodak sold the sales, marketing, and service operations to Danka, then the largest copier distributor in the world. In 1999 Kodak sold its copier manufacturing operations to Heidelberg. Since then Kodak-brand high-volume copiers have not had any significant market share in the United State. On the other hand, Kodak continues to be the dominant manufacturer of micrographics equipment and in 2000 purchased its main competitor, Bell & Howell.

Independent service organizations for photocopiers have largely disappeared in the United State. A 2000 federal circuit court opinion in a similar case against Xerox ruled ambiguously against ISOs when the OEM has patents on its parts (see "Postscript II" below). Further, since the mid-1990s

"This is not so far-fetched. Kodak argued at trial that the line should be drawn not just around its equipment and complementary service, but around every product Kodak produced that a service customer might buy. Tr. at 6010-6014. Thus, Kodak argued for a presumption of "systems" whenever there are complementary products and that Kodak's "reputation" is sufficiently important to make a wide range of superficially unrelated product effective complements: Kodak couldn't overcharge micrographic service customers because they might, someday, be a prospective buyer for Kodak copiers.
manufacturers have designed their machines so that technicians often must use copyrighted software in order to service them and have refused to provide the software to unauthorized ISOs. Lawsuits against this practice largely have been unsuccessful. It appears that the ISOs who brought suit against Kodak have either gone out of business or have become authorized distributors for other brands.

POSTSCRIPT I: DID THE SUPREME COURT LIMIT CLAIMS TO ONLY ONE ECONOMIC THEORY?

There have been numerous circuit court opinions on aftermarket cases since the Supreme Court (1992) decision in *Kodak,* and the results are deeply divided on a fundamental issue. At least three of these cases, we believe, clearly misread *Kodak.*

These three opinions turned on the same question: Was the allegation of *installed base opportunism* the crucial issue in *Kodak?* The First, Sixth, and Seventh Circuits have now each held that a *surprise* change in policy is necessary for a finding of a separate aftermarket that can be monopolized when the foremarket is competitive. Each has dismissed a plaintiff's claims, because there was no evidence or allegation of IBO.

The Supreme Court did not write that IBO was a single, special circumstance that permitted aftermarket power. The Court wrote:

> The fact that the equipment market imposes a restraint on prices in the aftermarket by no means disproves the existence of power in those markets. . . . Thus, contrary to Kodak's assertion, there is no immutable physical law—no "basic economic reality"—insisting that competition in the equipment market cannot coexist with market power in the aftermarket. *Kodak,* 504 U.S. at 471 [emphasis added]).

The Court emphasized that "marketplace realities" and market imperfections, such as high information costs and lock-in from switching costs, "could create a less responsive connection between service and parts prices and equipment sales" *Kodak,* 504 U.S. at 473 [emphasis added]).

The *Kodak* Court made specifically the point that we described in the theoretical section above: The key factor in whether or not aftermarket

---

5 The three we believe are in error are: *Lee v. Life Ins. Co. of North America,* 21 F.3d 14 (1st Cir. 1994), cert. denied, 513 U.S. 964 (1994); *PSI v. Honeywell,* 104 F.3d 811 (6th Cir. 1997), cert. denied (1997); and *Digital Equipment Corp. v. Uniq Digital Techs., Inc.,* 73 F.3d 756 (7th Cir. 1996), 520 U.S. 1265.

59 For example, the Sixth Circuit held "that an antitrust plaintiff cannot succeed on a Kodak-type theory when the defendant has not changed its policy after locking-in some of its customers." *PSI v. Honeywell,* 104 F.3d at 820.
power is possible is the strength of the link between the fore- and aftermarket responses to supracompetitive aftermarket pricing. The Court noted that every monopolist faces some constraints on its prices; the crucial question is the cross-elasticity between the aftermarket and the foremarket: How strong is the link? 

This is the key factual issue for market definition in all antitrust cases. For example, slide rules can, to a degree, substitute for computers: Is the linkage strong enough that consumers are protected from computer monopolization by slide rule competition? The Court agreed with the plaintiffs that the connection could be sufficiently weak due to IBO. The Court also stated that switching costs and high information costs could make the connection sufficiently weak. In short, the Court stated that one theory under which the systems view could fail is IBO; it did not state that this is the only theory. This point has been reinforced by the post-Kodak economic literature. Even one of Kodak’s own expert economists identified four different theories under which there is an opportunity to behave monopolistically in aftermarkets, with a resulting reduction in consumer welfare (Shapiro 1995).

There is a consistent economic theme here: Firms wish to maximize profits and thus want to act like monopolists given the opportunity. However, only sometimes are monopoly profits possible. The crucial economic question is what market constraints there are on a firm’s ability to charge supracompetitive prices or otherwise earn excess profits at the expense of consumers. When studying behavior in an aftermarket, one such market constraint may be the effect that aftermarket pricing has on foremarket profitability. This link should be analyzed in an aftermarket case. But there are any number of different circumstances under which the link is not sufficient to prevent monopoly harm. Simple, but incomplete economic theories should not provide antitrust immunity for entire classes of potentially harmful behavior. If plaintiffs in a case have a plausible alternative theory of aftermarket power and allege facts consistent with the theory, the case should proceed to trial, as did Kodak.

POSTSCRIPT II: DO INTELLECTUAL PROPERTY RIGHTS GRANT IMMUNITY FROM ANTITRUST?

After the Kodak case was completed, another issue that arose late in the proceedings has become very controversial, and another conflict has emerged

“We think the lower court’s opinion in Red Lion Medical Safety, Inc. v. Ohmeda, Inc. characterizes the economic issue properly: “The policy change [in Kodak] did not create the lock-in; instead, the existence of lock-in—high switching costs—made it both possible and economically desirable for Kodak to change its policy and exploit aftermarket consumers. Thus, the policy change did not create monopoly power; it was merely persuasive evidence that Kodak had market power in parts and engaged in monopolistic conduct in the aftermarket despite competition in the equipment market. To insist on a showing of a policy change confuses a symptom of market power and a lack of cross elasticity with the underlying condition itself (63 F. Supp. 2d 1218 at 1230 (E.D. Cal. 1999)).
between appeals courts. The controversy concerns the extent to which patents should have protected Kodak from antitrust liability.

Kodak argued on appeal that the patents that it had gave it the right to refuse to sell parts to ISOs. The Ninth Circuit ruled that protecting intellectual property is presumed to be a valid, pro-competitive business justification for refusing to sell or license. However, the court stated that this was a rebuttable presumption and concluded that plaintiffs successfully rebutted the Kodak’s patent defense by showing that only about sixty five of ten thousand parts were covered by patents and that the author of the Kodak parts policy testified that he didn’t give any thought to protecting Kodak’s intellectual property when crafting the policy.

The Federal Circuit, which is a special court created to hear appeals on cases in which patent issues are central to the case, has more recently ruled in the Xerox case that intellectual property owners have a much more general immunity from antitrust law. “The antitrust laws do not negate the patentee’s right to exclude others from patent property.” The Xerox case is remarkably similar to Kodak (it also involves a refusal to sell high-volume copier parts to ISOs), but the result was different: The court stated that firms could refuse to sell patented parts regardless of their motivation, in particular even if the motivation was to foreclose competition in a separate market such as service.

This conflict between two federal appeals courts is stark, and many authors have written about it (see, e.g., Katz and Safer [2002] and Boyle et al. [2002]). The issue highlights a long-standing area of controversy. Patent and other intellectual property laws establish property rights, which may create limited monopolies, in order to encourage invention and authorship and endow the property owners with the right to engage in exclusionary conduct. Both are seeking to enhance consumer welfare: intellectual property law by leading to new ideas and inventions, antitrust law by limiting allocative inefficiency (high prices, reduced quality or variety) by firms with current market power. Thus, we have a situation in which both types of law are largely

In re Independent Service Organizations Antitrust Litigation (Xerox), 203 F.3d 1322 at 1325 (Fed. Cir. 2000), cert. denied, 531 U.S. 1143 (2001), quoting Intergraph Corp. v. Intel Corp., 195 F.3d 1346 at 1362 (Fed. Cir. 1999). There is substantial uncertainty about what the Xerox decision even means. For example, in U.S. v. Microsoft (253 F.3d 34 at 63) the U.S. Court of Appeals for the District of Columbia cited Xerox when it wrote:

Microsoft’s primary copyright argument borders upon the frivolous. The company claims an absolute and unfettered right to use its intellectual property as it wishes: “If intellectual property rights have been lawfully acquired,” it says, then “their subsequent exercise cannot give rise to antitrust liability.” Appellant’s Opening Br. at 105. That is no more correct than the proposition that use of one’s personal property, such as a baseball bat, cannot give rise to tort liability. As the Federal Circuit succinctly stated: “Intellectual property rights do not confer a privilege to violate the antitrust laws.” In re Indep. Serv. Orgs. Antitrust Litig., 203 F.3d 1322, 1325 (Fed.Cir.2000).

Yet, Microsoft’s quoted argument above also claimed to rely on Xerox for support.
trying to accomplish the same thing (maximizing consumer welfare), but they come into conflict.

From an economic perspective, there should be some balancing: In some situations antitrust should restrict the way in which firms use their intellectual property, and in other situations intellectual property should be exempt from antitrust restrictions. The Supreme Court has repeatedly stated that intellectual property laws grant property rights, and those property rights are sometimes limited by antitrust (and other) laws. For example, in *Kodak* it wrote "[we have] held many times that power gained through some natural advantage such as a patent, copyright, or business acumen can give rise to liability if 'a seller exploits his dominant position in one market to expand his empire into the next.'" However, at present the law is very unclear about when antitrust limits the ability of firms to refuse to sell patented goods or to otherwise condition sale or licensing on anticompetitive conditions.

The Kodak case continues to be one of the most controversial antitrust cases of the modern era, although the controversy has expanded beyond whether aftermarket can be profitably monopolized to also include conflicts between intellectual property and antitrust law. The particular conflict between *Kodak* and *Xerox* almost surely will need to be resolved by the Supreme Court in some future case. Today, some prominent observers believe that 'recent cases, and particularly the Federal Circuit's opinion in *Xerox*, have upset that traditional balance [between intellectual property and antitrust] in a way that has disturbing implications for the future of antitrust in high-technology industries' (Pitofsky 2001). Other equally notable analysts believe that *Kodak* got it wrong, and *Xerox* got it right (Carlton 2001). Until the Supreme Court resolves the dispute, businesses and trial courts will not know what the law really is.

REFERENCES


