1-1-2000

Antitrust beyond Competition: Market Failures, Total Welfare, and the Challenge of Intramarket Second-Best Tradeoffs

Peter J. Hammer
Wayne State University

ANTITRUST BEYOND COMPETITION: MARKET FAILURES, TOTAL WELFARE, AND THE CHALLENGE OF INTRAMARKET SECOND-BEST TRADEOFFS

*Peter J. Hammer**

TABLE OF CONTENTS

* Assistant Professor, University of Michigan Law School. B.S. 1986, B.A. 1986, Gonzaga; J.D. 1990, Ph.D. 1993, University of Michigan. — Ed.

INTRODUCTION

Should antitrust law ever sanction the accumulation of market power or permit other restraints of trade if such conduct would increase social welfare? This is the challenge raised by intramarket second-best tradeoffs.[1] The lesson of second-best analysis is that one market failure can sometimes counteract the effects of another market failure. In the presence of multiple market failures, it is conceivable that mergers or other restraints traditionally viewed as anticompetitive may be welfare-enhancing. A social planner, given the mandate of maximizing total welfare, would permit such restraints. Could an antitrust judge come to the same result under a defensible application (or extension) of existing legal doctrine? This question highlights the tensions between an antitrust policy dedicated to preserving "competition" and an antitrust policy dedicated to maximizing total welfare. In doing so, the question tests the theoretical and practical limits of antitrust law,[2] asking whether it is time for antitrust law to move beyond structural understandings of competition and into the realm of express welfare analysis.[3]

---

1. This Article distinguishes intramarket second-best tradeoffs (tradeoffs involving multiple market failures in a single economic market) from intermarket second-best tradeoffs (tradeoffs involving multiple market failures in distinct economic markets). The proposed antitrust defense is limited to intramarket concerns largely for practical reasons. Intramarket concerns can be evaluated within a partial equilibrium economic framework, while many intermarket second-best tradeoffs cannot. The general theory of second best, however, teaches that intermarket second-best concerns can undermine intramarket second-best assessments. I defend the viability of antitrust law policing intramarket tradeoffs while ignoring intermarket tradeoffs on two grounds. First, a market definition process that strives to incorporate strong economic interrelationships should capture the second-best concerns of primary significance. Second, it is a defensible institutional division of labor to assign antitrust courts the task of policing intramarket tradeoffs, while assigning Congress the role of policing intermarket second-best concerns. For a further discussion of these justifications, see *infra* notes 15-18 and 187-190 and accompanying text.

2. The theoretical challenge is to antitrust law's goals and its operative methodology. The practical challenge is to the institutional capacity of courts, testing whether it is possible to devise a claim that could be adjudicated at acceptable administrative and error costs. As such, intramarket second-best tradeoffs test the limits of antitrust law and help reassess the proper relationship between antitrust law and antitrust economics. *See, e.g.*, Derek C. Bok, *Section 7 of the Clayton Act and the Merging of Law and Economics*, 74 HARV. L. REV. 226 (1960); Oliver E. Williamson, *Allocative Efficiency and the Limits of Antitrust*, 59 AM. ECON. REV. 105 (1969); Frank H. Easterbrook, *The Limits of Antitrust*, 63 TEXAS L. REV. 1 (1984).

3. "Competition" is a wonderfully ill-defined term — "anticompetitive" even more so. Most courts adopt some model or implicit understanding of what "competition" means and proceed to define "anticompetitive" as conduct that is inconsistent with their structural understanding. The weaknesses of this approach, and the comparative strengths of total welfare as an organizing principle for antitrust law, is one theme of this Article. I use the terms *total welfare*, *welfare-enhancing*, and *efficiency* to signify roughly interchangeable concepts.

This Article argues (1) that antitrust law should recognize a defense for private acts that restrain "competition" under the traditional antitrust analysis but advance total welfare, (2) that courts are competent to administer this defense, and (3) that the framework of existing antitrust statutes permits courts to recognize this defense. Today, most judges and scholars would reject intramarket second-best arguments as a justification for otherwise impermissible acts that enhance or maintain private market power, regardless of the credibility of the underlying economic analysis. The reasons given for rejecting such claims, however, would vary with the commentator. Some would argue that intramarket second-best claims have no statutory basis. Others would argue that maximizing total welfare is not the goal of the antitrust laws. Still others, who might be sympathetic to an efficiency-oriented antitrust doctrine, would be skeptical of the ability of the courts to implement a total welfare standard.

This Article challenges such received wisdom. Part I examines the theory of intramarket second-best analysis and explores how it fits within contemporary antitrust scholarship. Analytically, a defense for second-best tradeoffs has much in common with Oliver Williamson's productive efficiency defense in merger cases and with various market failure defenses that have been advanced in the literature. These similarities are examined, and the parameters of an affirmative defense are outlined.

I suggest that to rebut a finding of illegality based upon a traditional presumption of anticompetitive effects, defendants should have to establish (1) that the challenged conduct is responsive to an identifiable market failure; (2) that the conduct produces a net increase in total welfare (static efficiency); (3) that the conduct will not substantially impair subsequent efforts to address the underlying market failure (dynamic efficiency); and (4) that there is not a less restrictive course of action consistent with the antitrust laws that could achieve the same static efficiency gain. This defense presupposes a total welfare standard of analysis.

Part II examines the technical viability of the affirmative defense. Section II.A outlines the theoretical basis of consumer surplus and explores the tools underlying economic welfare analysis. Section II.B advances two contentions. First, the economic tools underlying many

---

In contrast, antitrust courts often speak of balancing procompetitive and anticompetitive effects. In some cases, these approaches overlap, and a court that is balancing pro- and anticompetitive effects is implicitly engaging in total welfare analysis. I argue that it would be a natural evolutionary step for courts to move beyond structural understandings and the rhetoric of procompetitive effects, and to engage expressly in welfare analysis. Moreover, taking this step need not involve a radical departure from existing doctrine, because second-best concerns and welfare analysis can be layered over existing competition-based antitrust presumptions, which are often effective proxies for economic efficiency. Differences will arise only when economic theory and empirical evidence call these presumptions into question.

intramarket second-best claims are sufficiently advanced to permit parties to explore such claims in the context of litigation. Second, the logic underlying a total welfare approach can be made intuitively accessible to judges, lawyers, and jurors. Indeed, a total welfare standard may actually provide better criteria for drafting jury instructions and counseling clients than an antitrust standard based on protecting "competition." Section II.C assesses the costs and benefits of implementing an intramarket second-best defense and examines how the doctrine can be defined to balance expected type one and type two errors.

Part III explores whether second-best tradeoffs and the total welfare standard underlying them can be reconciled with contemporary understandings of antitrust law and the institutional role of the courts. I argue that antitrust law is best understood in terms of an ongoing partnership between Congress and the courts. The reality of this partnership has important implications for statutory interpretation. While I disagree with scholars like Robert Bork, who claim that a total welfare standard is mandated by the antitrust statutes, I maintain that such an approach lies within the realm of acceptable evolutionary paths of judge-made antitrust doctrine. Consequently, the case for second-best analysis should be won or lost on the strength of its normative and policy justifications. I argue that a narrowly tailored affirmative defense restricted to intramarket tradeoffs creates an appropriate division of labor between Congress and the courts, and that a total welfare standard can be reconciled with the institutional role of the judiciary.

This Article presents only a skeletal outline of the second-best defense. Its purpose is to make it intellectually respectable to argue for the defense, not to demonstrate that in fact it should be allowed in any particular case. The examples that are provided suggest the kinds of seeming restraints that may deserve legitimation under this analysis, but to actually argue for these restraints in real cases would require more case-specific data and nuanced analysis. Only after courts and litigants have wrestled with second-best defenses in at least several real cases can the workability of the theoretical case I make here be fully evaluated. Broad abstract description, and an expression of faith in the combined capacities of courts and economists, only gets us started. Indeed, the best starting point may be to encourage enforcement officials to take second-best considerations into account in deciding whether to challenge particular restraints. This would permit the gradual accumulation of experience, which, if the decisions are incorporated in policy statements or made objects of study, may inform courts when they face intramarket second-best claims in litigation.

I. TOTAL WELFARE, MARKET FAILURES, AND INTRAMARKET SECOND-BEST CLAIMS: TOWARD DEFINING AN ANTITRUST DEFENSE

A. *Market Failures and the General Theory of Second Best*

According to the first theorem of welfare economics,[4] competitive equilibria are Pareto-efficient, meaning that there exists no reallocation of resources that could make someone better off without making someone else worse off.[5] This is a statement about the economy as a whole (general equilibrium), and envisions that when *all* industries and markets are competitive, the interaction between them will yield an efficient outcome. Economic proofs of the existence and efficiency of general competitive equilibria entail many restrictive assumptions. Some of these are: buyers and sellers act independently rather than collusively, resources are perfectly mobile and fungible, there are no production or consumption externalities, buyers know all relevant prices and qualities, and sellers know all production possibilities.[6] If *any* of these assumptions fail in *any* market, then the efficiency of the resulting general equilibrium (if one exists) can no longer be asserted.

If one condition fails in one market, then the policy prescription is straightforward: remedy the isolated failure, and the result will be a "first-best" solution. If there are multiple failures in a single market, or multiple failures in multiple markets, the analysis becomes more complicated. As Lipsey and Lancaster demonstrated in their analysis of the "general theory of second best," the efficiency of competitive equilibria is an all-or-nothing proposition.[7] Unless all conditions can

4. *See* HAL R. VARIAN, INTERMEDIATE MICROECONOMICS: A MODERN APPROACH § 29.7 (1987) (discussing the first welfare theorem).

5. For discussions of Pareto-efficiency, see DAVID M. KREPS, A COURSE IN MICROECONOMIC THEORY 153-56 (1990); VARIAN, *supra* note 4, §§ 17.9, 28.3; HAL R. VARIAN, MICROECONOMIC ANALYSIS 5, 198, 203 (2d ed. 1984).

6. For discussions of general equilibrium theory and its assumptions see KREPS, *supra* note 5, at 199-205; VARIAN, *supra* note 4, § 28.8, 28.12; VARIAN, *supra* note 5, at 189-211. *See generally* BRYAN ELLICKSON, COMPETITIVE EQUILIBRIUM: THEORY AND APPLICATIONS (1993). All economic modeling involves some level of abstraction and the making of assumptions. The assumptions need not be descriptively accurate for the model to produce meaningful insights. The failure of particular assumptions, however, can appropriately call the validity of the entire model into question.

7. *See* R.G. Lipsey & R.K. Lancaster, *The General Theory of Second Best*, 63 REV. ECON. STUD. 11, 11 (1956) ("The general theorem for the second best optimum states that if there is introduced into a general equilibrium system a constraint which prevents the attainment of one of the Paretian conditions, the other Paratian conditions, although still attainable, are, in general, no longer desirable."); *see also* C.E. FERGUSON, A MACROECONOMIC THEORY OF WORKABLE COMPETITION 15-17, 49-50 (1964) (discussing origins of second-best theory and implications of second-best considerations for theories of workable competition); RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW § 93, at 301 n.1 (5th ed. 1998) (providing an example of how second-best problems can undermine traditional efficiency analysis); CATHERINE M. PRICE, WELFARE ECONOMICS IN THEORY AND PRACTICE 31-42

be satisfied in all markets, there is no guarantee that remedying separate market failures will improve efficiency. Indeed, the counterintuitive proposition that remedying isolated market failures could actually make outcomes worse becomes possible:

> From this theorem there follows the important negative corollary that there is no *a priori* way to judge as between various situations in which some of the Paretian optimum conditions are fulfilled while others are not. Specifically, it is *not* true that a situation in which more, but not all, of the optimum conditions are fulfilled is necessarily, or is even likely to be, superior to a situation in which fewer are fulfilled.[8]

These insights can easily be leveraged into a full-scale attack on "piecemeal welfare economics."[9] Richard Markovits has explored many of the implications of second-best theory for law and economics.[10] In particular, Markovits has criticized courts and scholars for failing to appreciate the implications of second-best theory for contemporary antitrust doctrine.[11]

Antitrust scholars have reacted to the general theory of second best either by dismissing its implications entirely, or by using its insights to invalidate economic approaches to antitrust altogether. Both responses are misguided. Many economically oriented scholars acknowledge second-best problems but nevertheless reject their implications. A range of justifications is provided: some contend that incorporating second-best concerns would be too complicated for courts to handle, others argue that a simple heuristic of promoting competition on a serial basis will lead to the most defensible results, and others

---

(1977) (providing a comprehensive discussion of the theory of second best); MICHAEL WATERSON, ECONOMIC THEORY OF INDUSTRY 5 (1984) (discussing the general theory of second best and its implications for welfare economics); C.G. Veljanovski, *The New Law-and-Economics: A Research Review*, *in* READINGS IN THE ECONOMICS OF LAW AND REGULATION 21-22 (A.I. Ogus & C.G. Veljanovski eds., 1984) (same).

8. Lipsey & Lancaster, *supra* note 7, at 11-12.

9. *Id.* at 17.

10. *See, e.g.*, Richard S. Markovits, *The Allocative Efficiency of Shifting from a "Negligence" System to a "Strict Liability" Regime in Our Highly-Pareto-Imperfect Economy: A Partial and Preliminary Third-Best Allocative-Efficiency Analysis*, 73 CHI.-KENT L. REV. 11 (1998); Richard S. Markovits, *A Basic Structure for Microeconomic Policy Analysis in Our Worse-Than-Second-Best World: A Proposal and Related Critique of the Chicago Approach to the Study of Law and Economics*, 1975 WIS. L. REV. 950; Richard S. Markovits, *The Limits to Simplifying Antitrust: A Reply to Professor Easterbrook*, 63 TEXAS L. REV. 41 (1984) [hereinafter Markovits, *The Limits to Simplifying Antitrust*]; Richard S. Markovits, *Monopoly and the Allocative Inefficiency of First-Best-Allocatively-Efficient Tort Law in Our Worse-Than-Second-Best World: The Whys and Some Therefores*, 46 CASE W. RES. L. REV. 313 (1996); Richard S. Markovits, *Monopolistic Competition, Second Best, and the Antitrust Paradox: A Review Article*, 77 MICH. L. REV. 567 (1979) [hereinafter Markovits, *Second Best, and the Antitrust Paradox*]; Richard S. Markovits, *Second-Best Theory and Law & Economics: An Introduction*, 73 CHI.-KENT L. REV. 3 (1998).

11. *See* Markovits, *The Limits to Simplifying Antitrust*, *supra* note 10, at 45-48; Markovits, *Second Best, and the Antitrust Paradox*, *supra* note 10, at 577-80.

maintain that second-best problems must be ignored because Congress has established a legislative policy favoring competition.[12] Scholars who are less sympathetic to an economically oriented antitrust law have employed second-best theory to undermine efficiency as a normative antitrust benchmark and to bolster the case for fairness or distributive justice as alternative guiding principles for antitrust law.[13]

We can better understand these issues if we distinguish general from partial equilibrium economic analysis, and intermarket from intramarket second-best concerns. General equilibrium analysis examines the simultaneous interaction of all markets in the economy. Maintaining an antitrust policy premised on such a framework is not practical. In this respect, the received antitrust wisdom is correct — considering the effects of market distortions elsewhere in the economy (intermarket second-best concerns) requires such wide-ranging knowledge of conditions throughout the economy and the way they would respond to changes in particular sectors, that parties could not accurately canvass them nor courts evaluate the complex implications of granting relief. Entirely different questions are raised, however, if we ask whether antitrust law should incorporate the effects of intramarket second-best claims in the context of partial equilibrium analysis. In their haste to reject intermarket second-best concerns, antitrust scholars have failed to appreciate the potential significance of intramarket second-best concerns.[14]

---

12. *See* ROBERT H. BORK, THE ANTITRUST PARADOX: A POLICY AT WAR WITH ITSELF 113-14 (1993) (rejecting second-best claims because of the inability of the judiciary to engage in the necessary economic analysis and because of a legislative policy in favor of competition); HERBERT HOVENKAMP, FEDERAL ANTITRUST POLICY: THE LAW OF COMPETITION AND ITS PRACTICE 38-39 (1994) (rejecting the general theory of second best as being "extraordinarily complex" and impractical); Herbert Hovenkamp, *Antitrust After Chicago*, 84 MICH. L. REV. 213, 241-42 (1985) ("Problems of second-best may be so overwhelming and so hypothetical that the antitrust policy maker is well off to avoid them."); *see also* 2 PHILLIP AREEDA & DONALD F. TURNER, ANTITRUST LAW 308-13 (1978) (acknowledging second-best concerns both between and within markets, but advocating a general policy of promoting competition on a serial basis).

13. *See* Robert G. Harris & Thomas M. Jorde, *Antitrust Market Definition: An Integrated Approach*, 72 CAL. L. REV. 1, 13-14 (1984) (arguing that second-best concerns undermine efficiency, but not distributive or fairness justifications for antitrust law); Lawrence A. Sullivan, Book Review, 75 COLUM. L. REV. 1214, 1219-20 (1975) (arguing that the general theory of second best undermines allocative efficiency as a normative guide for antitrust law).

14. *See* Peter J. Hammer, Mergers, Market Power and Competition: An Economic and Legal Evaluation of Hospital Mergers, at 290 n.137 (1993) (unpublished Ph.D. dissertation, University of Michigan, Dep't of Econ.) (on file with author) ("To the extent that these conclusions pertain to welfare implications between distinct markets from a standpoint of general equilibrium theory, the literature [advocating the rejection of second-best claims] is correct. To the extent that the second best concerns are raised in the same economic market, the claims may deserve more serious attention than they have received to date.").

Partial equilibrium analysis seeks to examine a single well-defined market in isolation, holding the conditions in other markets constant.[15] Here, data problems are often tractible, and current economic tools provide a means of addressing legitimate second-best concerns while avoiding the paralysis inherent in the general theory of second best. Oliver Williamson, responding to Lawrence Sullivan's second-best attack on the ability of antitrust courts to rely upon economic principles,[16] has defended the legitimacy of employing partial equilibrium analysis in antitrust disputes.[17] Williamson argues that a properly constructed partial equilibrium framework can partition the economic analysis in a manner that captures the most significant second-best effects, while properly ignoring remote second-best concerns. "Strong interaction effects then can be taken expressly into account, and elsewhere the second-best qualification deserves the weight that lawyers label 'de minimis.' "[18]

The arguments Williamson gives to justify partial equilibrium analysis in antitrust cases militate in favor of expanding antitrust analysis to accommodate intramarket second-best concerns. Strong economic interconnections not only can be taken into account, they should be incorporated into the economic analysis. The theoretical

---

15. *See* KREPS, *supra* note 5, at 279-80 (discussing limitations of partial equilibrium analysis); POSNER, *supra* note 7, § 3.2, at 93 (contrasting partial and general equilibrium analysis); VARIAN, *supra* note 4, at 480 (same); VARIAN, *supra* note 5, at 3-5 (introducing concept of equilibrium and the framework of partial equilibrium analysis). Obviously, conditions in all other markets will not remain constant. Changes in one market can have ripple effects in others. Similarly, changes elsewhere in the economy can influence the market being examined. If these economic interconnections are strong, the partial equilibrium framework should be defined to incorporate them into the analysis. If the interconnections are not strong, then one can often ignore such second-order effects without undermining the validity of the analysis.

16. *See* Sullivan, *supra* note 13, at 1220 ("Absent the simultaneous fulfillment of all conditions of optimum allocation ... economic theory tells us nothing about how to improve resource allocation.... Economics simply provides no basis on which to say, for example, that ending monopoly in the shoe industry or ending a price cartel in the electrical equipment industry (or doing both of these things) will improve resource allocations and increase aggregate welfare. Given the persistence of other deviations (other monopolies, cartels, tariffs, and distorting taxes), there is no basis for assuming that doing away with any one or more deviations from optimality will improve efficiency at all.").

17. *See* Oliver E. Williamson, *Assessing Vertical Market Restraints: Antitrust Ramifications of the Transaction Cost Approach*, 127 U. PA. L. REV. 953, 986-88 (1977) (rejecting Sullivan's critique of antitrust efficiency analysis and defending the ability of partial equilibrium analysis to evaluate antitrust issues).

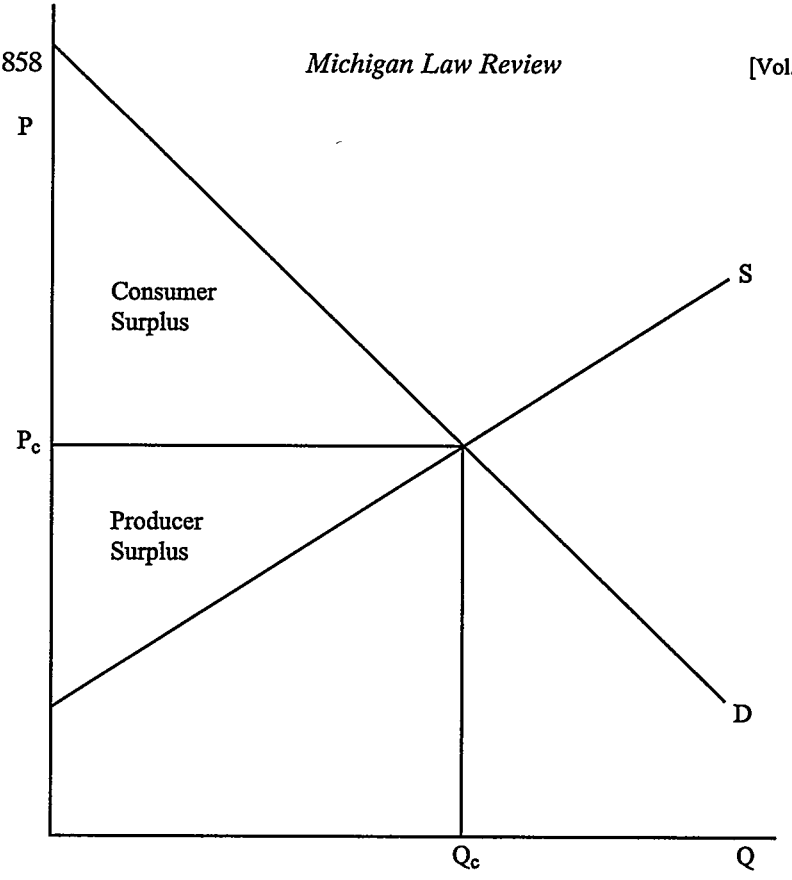
18. *Id.* at 987. There is an unstated assumption in Williamson's argument. A properly framed partial equilibrium analysis may be able to partition the economy in a manner that accommodates strong economic interconnections, while ignoring remote second-best consequences, but the economist's concept of "partitioning" the economy must be mapped onto the antitrust concept of defining the relevant market. To bridge this gap, courts should attempt to define antitrust markets in a manner that corresponds with a defensible partial equilibrium analysis. If they do so, then it is meaningful to distinguish between intramarket and intermarket second-best problems, and to construct an antitrust defense that addresses the former while defensibly ignoring the latter.

and practical justifications relied upon to reject general second-best theory and intermarket second-best concerns do not apply to intramarket second-best problems. Indeed, whether they acknowledge it or not, antitrust courts and scholars engage in partial equilibrium analysis all the time, implicitly ignoring the general second-best implications of their analysis. From this perspective, my proposal is relatively uncontroversial: the same tools that antitrust courts currently use to examine a full range of economic activity should also be used to examine a range of intramarket second-best concerns as well.[19]

What then does partial equilibrium analysis look like and what does it seek to accomplish? Within this economic framework, the judge or decisionmaker is concerned about the effect that market changes will have on total welfare, with total welfare defined as the sum of producer and consumer surplus (see figure 1).[20]

19. The dilemma is actually fairly sharp. Either the tools of partial equilibrium analysis are robust enough to address ordinary antitrust disputes, in which case they are robust enough to address the implications of intramarket second-best tradeoffs, or the economic tools of partial equilibrium analysis are not robust enough to assess the implications of intramarket second-best tradeoffs, in which case their ability to assess the economic consequences of a wide range of traditional antitrust disputes must also be called into question. Obviously, one can raise concerns about the ability of courts to implement the theory in practice, or whether such an analysis is consistent with the antitrust statutes, issues that are addressed *infra* Parts II and III, respectively; but so long as one is willing to conceed the propriety of adopting a partial equilibrium framework in antitrust cases, it is difficult to simply ignore the problems raised by intramarket second-best concerns.

20. Simplistically, profits can be thought of as the area between the upward sloping supply curve and the market price for the product (see figure 1). The concept of consumer surplus can be made graphically concrete by envisioning the area between the downward sloping market demand curve for Product A and the price for Product A (see figure 1). A technically more accurate definition of consumer surplus is provided *infra* notes 111-114 and accompanying text. For a basic introduction to the concept of consumer and producer surplus, see HOVENKAMP, *supra* note 12, at 4-5, and VARIAN, *supra* note 4, §§ 15.1, 15.11.



Producer surplus is a measure of producer profits. Consumer surplus is the aggregation of the differences between all consumers' individual reservation prices (how much they would be willing to pay for Product A) and the market price (how much they have to pay for Product A),[21] where "consumers" are defined as actual purchasers of Product A and those who would have purchased Product A if the price had been lower.[22] Conceptually, a rational antitrust policy could

21. This Article employs the terminology of traditional economic analysis: total welfare, consumer surplus, and producer surplus. Substantial confusion exists in the antitrust literature because of Judge Bork's idiosyncratic use of the term "consumer welfare." *See* Robert H. Bork, *Legislative Intent and the Policy of the Sherman Act*, 9 J.L. & ECON. 7, 7-16 (1966) (defining consumer welfare implicitly in terms of total welfare); BORK, *supra* note 12, at 90-91 (defining consumer welfare in terms of maximizing productive and allocative efficiency — total welfare). Joseph Brodley has rightly declared that "consumer welfare is the most abused term in modern welfare analysis." Joseph F. Brodley, *The Economic Goals of Antitrust: Efficiency, Consumer Welfare, and Technological Progress*, 62 N.Y.U. L. REV. 1020, 1032 (1987). The economist's "total welfare" and Judge Bork's "consumer welfare" are functionally equivalent. *See* John R. Morris, *International Trade and Antitrust: Comments*, 61 U. CIN. L. REV. 945, 945-46 n.4 (1993). Although I disagree with Bork's choice of words, I am generally sympathetic to the contention that antitrust law should strive to maximize total welfare.

22. This common sense understanding of consumers contrasts sharply with that of Judge Bork's. Bork defines "consumers" to include the owners of firms and monopolies. *See* BORK, *supra* note 12, at 108-10. Understandably, if "producers" are also "consumers," then

be designed to maximize either total welfare or consumer surplus.[23] The choice of which to maximize, as we shall see, can have important ramifications for antitrust policy.

The basic tools of partial equilibrium analysis provide a reasonable framework in which to examine the effects of intramarket second-best tradeoffs. Still, a number of important questions persist. While the theoretical and pragmatic arguments antitrust scholars give to justify ignoring intermarket second-best claims are not persuasive when applied to intramarket second-best claims, can a theory of intramarket second best be designed that courts are capable of implementing? Does the antitrust laws' commitment to "competition" preclude considering intramarket second-best tradeoffs in cases where the goals of competition and total welfare conflict? Before attempting to answer these questions, it is useful to gain a better understanding of exactly what an analysis of intramarket second-best tradeoffs would entail.

B. *Illustrations of Intramarket Second-Best Tradeoffs*

A variety of intramarket second-best tradeoffs are possible. I begin with the example of allowing a concentration of market power to counteract negative externalities. Although problems of negative externalities might better be addressed by taxes or regulation, they are an easily understood example of intramarket second-best tradeoffs. I

"consumer welfare" must be equal the sum of producer and consumer surpluses (the economist's definition of total welfare).

23. Robert Lande is the leading proponent of the consumer surplus standard, although he phrases his theory in terms of avoiding "wealth transfers." *See* Robert H. Lande, *Wealth Transfers as the Original and Primary Concern of Antitrust: The Efficiency Interpretation Challenged*, 34 HASTINGS L.J. 65 (1982) [hereinafter Lande, *Wealth Transfers*]. A standard preventing wealth transfers creates a property right, giving consumers an entitlement to the difference between their reservation price and the competitive market price. *See* Alan A. Fisher et al., *Price Effects of Horizontal Mergers*, 77 CAL. L. REV. 777, 787 n.31 (1989); Robert H. Lande, *Commentary: Implications of Professor Scherer's Research for the Future of Antitrust*, 29 WASHBURN L.J. 256, 260 (1990). Not surprisingly, a total welfare and a consumer surplus standard will often produce different antitrust results. *See* Peter J. Hammer, *Questioning Traditional Antitrust Presumptions: Price and Non-Price Competition in Hospital Markets*, 33 MICH. J.L. REF. (forthcoming 1999) (classifying mergers in terms of their differential effects on total, producer, and consumer surpluses and exploring the difference between a total welfare and a consumer surplus standard). For example, a consumer surplus standard would substantially redefine the contours of Oliver Williamson's productive efficiency defense. *See infra* Section I.C.1. Increases in productive efficiency (cost reductions) could still be a defense to a merger, but the cost reductions must be so large that they yield lower post-merger prices. *See* Alan A. Fisher & Robert H. Lande, *Efficiency Considerations in Merger Enforcement*, 71 CAL. L. REV. 1580, 1630 (1983). Under this approach, gains in productive efficiency that would increase total welfare and save social resources, where the savings were manifested in increased producer profits rather than lower prices, would not be permitted. Similarly, a consumer surplus rule would not permit intramarket second-best tradeoffs under the Kaldor-Hicks standard. *See* discussion *infra* note 46 and accompanying text. The fact that producers gained more than what consumers lost would not be a sufficient warrant for policy action. Any reduction in consumer surplus would constitute a sufficient basis for rejecting the challenged practice.

then present the more plausible case of hospital nonprice competition, where a merger might curb wasteful forms of competition, such as providing private parking spaces to physicians or excessive duplication of services and technology. Finally, I offer examples that illustrate the breadth of arrangements that may raise second-best concerns.

1. *Negative Externalities*[24]

Markets are imperfect and often fail.[25] In the presence of market failures, arrangements that might otherwise be considered undesirable, like the presence of a monopoly, may in fact be welfare-enhancing. The pricing rule characteristic of competitive markets equates price with private marginal costs. When a negative externality, like pollution, exists, private actors fail to internalize the full social costs of their actions.[26] As a result, the competitive price is lower and the competitive output is higher than is socially desirable.[27] The magnitude of the resulting allocative inefficiency can be measured by the shaded triangle in figure 2.

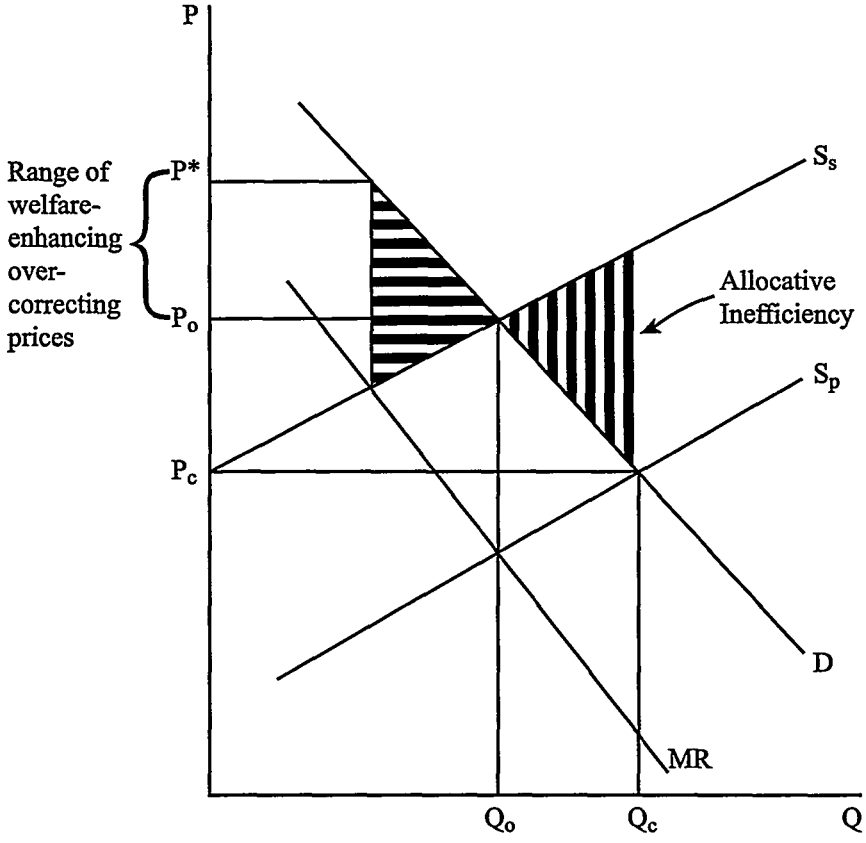
24. Negative externalities represent a lightning rod example that is used primarily for illustrative purposes. This is not the strongest case in favor of intramarket second-best analysis in antitrust law, nor the most likely application. The example of externalities, however, is both provocative and easy to grasp. The illustration also accurately suggests the ramifications (good and bad) of permitting antitrust courts to conduct intramarket second-best tradeoffs. *But see* Christopher R. Leslie, *Achieving Efficiency Through Collusion: A Market Failure Defense to Horizontal Price Fixing*, 81 CAL. L. REV. 243, 275-77 (1993) (advocating the adoption of a market failure defense for negative price externalities).

25. *See* ROBERT COOTER & THOMAS ULEN, LAW AND ECONOMICS 45-49 (1988) (discussing market failures such as monopoly power, externalities, public goods, and imperfect information).

26. *See id.* at 169-70 (discussing externalities and possible remedies); PAUL R. FERGUSON & GLENYS J. FERGUSON, INDUSTRIAL ECONOMICS 142-44 & fig.7.2 (2d ed. 1994) (discussing externalities and illustrations of the potential welfare loss); VARIAN, *supra* note 5, at 259-63 (discussing externalities and the range of possible solutions).

27. *See* FERGUSON & FERGUSON, *supra* note 26, at 143 fig.7.2 (illustrating the effects of negative externalities on price and output).

FIGURE TWO: NEGATIVE PRICE EXTERNALITIES



For example, assume that two private companies that are polluting the air merge to attain a monopoly. Their monopoly power allows them to raise the price of their goods, demand falls, production drops, and the air becomes cleaner, perhaps at a savings to society that exceed the social costs of the merger. True, a social problem that might be better curtailed by regulation than by allowing a monopoly does not make the most appealing case for efficiency-enhancing second-best tradeoffs, but it does illustrate how these can occur. Moreover, externalities or other inefficiencies are not always obvious targets for regulatory control, or circumstances may cause regulatory control to appear less promising than market solutions. Finally, antitrust courts, which must decide the cases presented to them, may only be able to increase efficiency by approving such mergers, and doing so does not preclude later regulation.

According to second-best theory, one market failure can sometimes counteract the effects of another market failure.[28] We have just seen how one market failure, market power, may, in the face of significant negative externalities, have such a countervailing effect. It follows that enforcing the ordinary antitrust prescription may produce an inefficient outcome,[29] while allowing the exercise of market power (either through merger or cartelization) may improve total welfare. Market power does not, however, ensure that other market failures, including those involving significant externalities, will be alleviated at all, much less to an extent that offsets the efficiency costs of the exercise of market power. The monopolist's rule of equating marginal revenue with private marginal costs could lead to an overcorrection that produces higher prices and lower outputs than are socially optimal. Importantly, however, even assuming overcorrection, there is a range of monopoly pricing where the allocative inefficiency associated with the monopoly pricing is less than the inefficiency associated with the negative externality (see figure 2). Any monopoly outcome within this range would yield an increase in total welfare, as compared to the competitive market. If antitrust law were guided by a total welfare standard, then merger or cartelization in this area would not be ruled out.[30]

Cigarette smoking provides an illustration.[31] Cigarette smoking imposes a number of external costs. To the extent that health care costs are paid through the Medicare or Medicaid program, a smoker fails to internalize all of the economic consequences of smoking-related illnesses. Smokers similarly fail to internalize the harmful effects of second-hand smoke. In a competitive market for cigarettes,

28. See discussion of the general theory of second best *supra* Section I.A.

29. *See* John Shepard Wiley, Jr., *A Capture Theory of Antitrust Federalism*, 99 HARV. L. REV. 713, 751-52 (1986) ("An established and abundant literature has elaborated the conditions under which competition enforced by ordinary Sherman Act rules produces inefficiency.").

30. Depending upon how the consumer's price entitlement is defined, a consumer surplus standard would not permit these intramarket second-best tradeoffs. Consumer surplus (at least as defined by the benchmark where the competitive price equals private marginal cost) would necessarily be reduced in moving from the inefficient competitive level of price and quantity to any more efficient allocation of resources. If the consumer surplus "entitlement" were defined at the socially efficient "price equals marginal social cost" level, then a consumer surplus standard would permit intramarket second-best tradeoffs in any range increasing price from P^C to P^O in figure 2, but would not permit any overcorrection by the would-be monopolist in the P^O to P^* range — even though there would be a net increase in total welfare in this range as compared to the competitive equilibrium.

31. *See* Ellen Gulbrandsen & Susan Skeath, Would Big Tobacco Have Been Better?: The Social Welfare Implications of Antitrust Action in the Presence of Negative Externalities, (Wellesley College Working Paper 99-07) (June 1999) (on file with author) (examining the welfare effects of the 1911 antitrust breakup of the American Tobacco Company and arguing that a monopoly market structure may have resulted in a higher level of social welfare than was produced by the breakup and the introduction of competition).

price would be equated to private marginal cost, leading to the overproduction and overconsumption of cigarettes; that is, there is more smoking than would exist in a world where all the social costs of the activity were internalized. In this setting, the exercise of private market power (either through merger or cartelization) would lead to an increase in price and a reduction in consumption. That might well increase social welfare.[32]

This does not mean that mergers of cigarette companies should necessarily be fostered. Comparative institutional analysis must play a role in implementing the theory of second best.[33] Rather than sanctioning private market power, the traditional policy response is to internalize the effects of the externality either through a tax[34] or through

32. One can imagine many factors that could complicate this simple analysis. The resources diverted from the cigarette industry might be used for production and consumption in a different market that itself may produce pollution or other social harms. The profits earned by the cigarette monopolist might be directed towards increased advertising and promotion to further boost cigarette demand and sales. Reduced levels of production may force small tobacco farmers out of work, or may lead to layoffs of employees at cigarette factories. Finally, many may think that smoking is inherently objectionable and that smoking should not be tolerated at all. Some of these concerns are implicitly addressed in the economic analysis. If the market is moving from one equilibrium to another, one would not expect there to be significant gains to promotional advertising that would not have previously been exploited. If the supply curve accurately reflects the social cost of resources, then directing resources out of the cigarette industry (a market with negative externalities) into other markets should produce an increase in economic welfare, even if production in alternative markets results in some negative effects. This being said, the existence of externalities inside markets and the existence of failures across markets — the very lesson of the general theory of second best — can directly undermine this assertion. In the end, all an economist can do is acknowledge these limitations and appeal to the fact that modeling necessarily involves some level of abstraction, *see supra* note 6, that proper partial equilibrium analysis strives to incorporate strong economic interrelationships, *see supra* note 15, and that ultimately the persuasiveness of the model's conclusions and the economic argument depends upon the context of its application and the seriousness of the particular objections to its assumptions.

Other social values, such as the interests of small tobacco farmers, the interests of cigarette factory workers, or the inherent social objections to smoking are not easily reflected in the economic analysis. This highlights the fact that economic modeling does not obviate the need to consider the appropriate relationship between economic and non-economic concerns in the political process.

33. The existence of federal and state cigarette taxes and potential tort liability complicates this simple example. Taxes can be used instead of an antitrust action, and tort liability can be used instead of or in addition to taxation. The same economic analysis used to conduct second-best analysis can be used to estimate the optimal level of taxation. Interestingly, if one looks at the cigarette industry today, the evidence suggests that taxes are substantially lower than what would be required to internalize the full social costs of smoking. *See* Gulbrandsen & Skeath, *supra* note 31, at 21-22. If this is true, then a story of the failure of the political process must be told at the same time as the story of the failure of economic markets. Given the reality of political market failures and the inability to impose appropriate taxes, there may be a role for antitrust second-best tradeoffs even in the realm of externalities.

34. *See* VARIAN, *supra* note 4, § 31.4 (illustrating the ability of Pigouvian taxes to internalize the effects of externalities and produce socially optimum outcomes).

the tort system.[35] An appropriate tax imposed upon a perfectly competitive market could produce the socially efficient outcome without creating private market power and rewarding businesses that some would regard as evil. Which approach is superior depends upon the relative administrative and error costs associated with the different institutional responses. The choice, moreover, need not be dichotomous. One can envision concurrent or competing systems of private second-best solutions and public regulation. Permitting the exercise of private market power as a second-best remedy will be problematic only if the solidification of market power would impair the ability of markets to correct themselves subsequently, or would impede the implementation of superior public solutions. In this sense, intramarket second-best analysis must be sensitive to dynamic as well as static efficiency concerns.[36]

2. *Hospital Nonprice Competition*

Competition in medical markets raises more complicated second-best claims. In traditionally structured health care markets with pervasive insurance coverage and passive third-party payors,[37] hospital competition emerged along nonprice rather than price dimensions.[38] The welfare effects of nonprice competition are ambiguous. Nonprice competition has been condemned by some as a medical arms race,[39] and praised by others for creating incentives to improve quality and provide better care to patients, for example, new technology, better doctors, improved facilities.[40]

---

35. *See* COOTER & ULEN, *supra* note 25, at 54-66 (examining the ability of negligence and strict liability tort rules to internalize the costs of accidents and lead to efficient levels of care).

36. See *infra* Section I.D for a discussion of dynamic efficiency concerns within the context of a proposed affirmative defense for intramarket second-best tradeoffs. This aspect of the defense examines whether the private "anticompetitive" conduct would increase the costs of subsequent public or private remedies.

37. *See* Hammer, *supra* note 23 (examining the economic structure of traditional health care markets).

38. *See id.* (exploring the incentives motivating hospital price and nonprice competition); Peter J. Hammer, *Price and Quality Competition in Health Care Markets: The Comparative Institutional Case Against an Antitrust Exemption for Medical Self Regulation*, *in* ACHIEVING QUALITY IN MANAGED CARE: THE ROLE OF LAW 131-34 (John D. Blum ed., ABA monograph 1997) (same).

39. *See, e.g.*, David Dranove et al., *Is Hospital Competition Wasteful?*, 23 RAND J. ECON. 247, 257-60 (1992) (discussing medical arms race scenario).

40. *See, e.g.*, American Med. Int'l, 104 F.T.C. 1, 185 (1984) ("Because . . . the nonprice competition that does exist in the industry responds to consumers' expressions of their wants by providing services valued by physicians and patients, such as expanding the range of choices available to them, stimulating innovation, and developing expertise by hospitals, we conclude that 'rivalry' among hospitals along nonprice dimensions constitutes competition in the economic sense that warrants protection by the antitrust laws."); United States v. Rockford Mem'l Corp., 717 F. Supp. 1251, 1283-84 (N.D. Ill. 1989) ("For similar reasons,

Models of hospital competition (similar to models of nonprice competition in the price-regulated airlines industry) demonstrate that nonprice competition is motivated both by a desire to stimulate aggregate market demand and by a desire to steal demand away from competitors.[41] To the extent that nonprice investments are motivated by consumer preferences embodied in the market demand function, such investments are likely to be cost-justified, increase social welfare, and be positively affected by market competition. To the extent that nonprice competition is motivated by its ability to steal rather than stimulate demand, such efforts are suspect from a welfare perspective.[42] While nonprice investments yield positive benefits to consumers, nonprice attributes may be supplied beyond the point where the marginal cost of such services equals their marginal benefits. Put differently, while consumers value nonprice amenities, they do not necessarily value them at a level where they would be willing to pay the full cost of the amenity if they had a choice.[43] When this is so, the social cost of providing the nonprice amenity will exceed the social benefits derived from its provision. In these circumstances, restricting nonprice competition increases total welfare.

This theoretical analysis comports with a substantial body of empirical evidence demonstrating a positive correlation between hospital competition and hospital expenses.[44] In traditionally structured hospi-

---

nonprice competition such as advertising and quality of care has also intensified, as each hospital desires to offer physicians and patients the latest medical technology and services in an effort to present itself as a full service high quality institution. The hoped-for result of these efforts, of course, is the attraction of quality physicians and more and better paying inpatients." (citations to record omitted)), *aff'd*, 898 F.2d 1278 (7th Cir. 1990).

41. *See* Philip J. Held & Mark V. Pauly, *Competition and Efficiency in the End Stage Renal Disease Program*, 2 J. HEALTH ECON. 95, 100-03 (1983) (modeling competition in health care markets); Hammer, *supra* note 14, at 103-18 (same). The dynamics of hospital nonprice rivalry in the Old Regime are structurally similar to the rivalry characteristic of the price-regulated airline industry in the 1970s. *See generally* Lawrence J. White, *Quality Competition and Regulation: Evidence from the Airline Industry*, *in* REGULATING THE PRODUCT (R.E. Caves & M.J. Roberts eds., 1975); Lawrence J. White, *Quality Variation When Prices Are Regulated*, 3 BELL J. ECON. 425 (1972) (discussing rivalry in the price-regulated airline industry).

42. *See* Hammer, *supra* note 23 ("Hospitals are concerned with the determinants of hospital-specific demand. Quality investments stimulate hospital-specific demand both by increasing market demand and, more importantly, by diverting demand from other hospitals. The 'demand-stealing' effects provide incentives to supply quantities of the nonprice attribute that may exceed the level justified by consumer preferences as reflected in the market demand function."). One economist has gone so far as to model the demand-stealing aspect of hospital nonprice competition as a prisoner's dilemma. *See* Joseph M. Jadlow, *Hospital Competition and the Prisoner's Dilemma*, 25 RIVISTA INTRERNAZIONALE DI SCIENZE ECONOMICHE E COMERCIALI 360 (1978).

43. *See* Hammer, *supra* note 23.

44. *See generally, e.g.*, MONICA NOETHER, COMPETITION AMONG HOSPITALS, (Staff Report of the Bureau of Economics, Federal Trade Commission) (1987); Held & Pauly, *supra* note 41, at 95; Paul L. Joskow, *The Effects of Competition and Regulation on Hospital Bed Supply and the Reservation Quality of the Hospital*, 11 BELL J. ECON. 421 (1980);

tal markets, higher levels of competition are associated with higher medical costs. Conversely, economic concentration is associated with correspondingly lower medical costs. This surprising phenomenon presents a difficult dilemma for policymakers. Market power in the hospital industry is likely to lead to both lower costs and, to the extent that nonprice competition is associated with quality, lower quality of care.[45]

The intramarket second-best tradeoffs in health care parallel the discussion of negative price externalities. Given the multiple failures in medical markets — imperfect information, moral hazard, agency problems — competition may lead to a level of nonprice rivalry that exceeds what is socially desirable, driving a wedge between the socially optimum equilibrium and the competitive equilibrium. In this setting, one market failure (market power or cartelization) may counteract the effects of other market failures (imperfect information, agency problems, and moral hazard). Consequently, it is at least plausible that mergers may lead to the provision of more efficient levels of nonprice services and a more rational allocation of medical resources.

We can phrase this welfare analysis as a Kaldor-Hicks compensation problem that illustrates the nature of the underlying tradeoff.[46] Consumers benefit from nonprice competition. The reduction in non-

---

Monica Noether, *Competition Among Hospitals*, 7 J. HEALTH ECON. 259 (1988); James C. Robinson & Harold S. Luft, *Competition and the Cost of Medical Care, 1972 to 1982*, 257 J. AM. MED. ASS'N 3241 (1987); James C. Robinson & Harold S. Luft, *The Impact of Hospital Market Structure on Patient Volume, Average Length of Stay, and the Cost of Care*, 4 J. HEALTH ECON. 333 (1985); George W. Wilson & Joseph M. Jadlow, *Competition, Profit Incentives, and Technical Efficiency in the Provision of Nuclear Medicine Services*, 13 BELL J. ECON. 472 (1982). For reviews of this literature, see Hammer, *supra* note 14, at 22-27; Frederic J. Entin et al., *Hospital Collaboration: The Need for an Appropriate Antitrust Policy*, 29 WAKE FOREST L. REV. 107, 153-67 (1994); and Paul A. Pautler & Michael G. Vita, *Hospital Market Structure, Hospital Competition, and Consumer Welfare: What Can the Evidence Tell Us?*, 10 J. CONTEMP. HEALTH L. & POL'Y 117, 123-29 (1994).

45. *See* Hammer, *supra* note 23; Peter J. Hammer, *Medical Antitrust Reform: Arrow, Coase and the Changing Structure of the Firm*, *in* THE PRIVATIZATION OF HEALTH CARE REFORM (Gregg Bloche ed., Oxford University Press, forthcoming 2000).

46. *See* Hammer, *supra* note 23 (discussing the welfare effects of hospital nonprice competition in terms of Kaldor-Hicks tradeoffs). For a discussion of the Kaldor-Hicks standard, see J.R. Hicks, *The Valuation of the Social Income*, 7 ECONOMICA 105 (1940), and Nicolas Kaldor, *Welfare Propositions of Economics and Interpersonal Comparisons of Utility*, 49 ECON. J. 549 (1939). *See also* Hovenkamp, *supra* note 12, at 239-40 (discussing the Kaldor-Hicks standard).

In this Article, I embrace forms of welfare analysis that go beyond simple Pareto-efficiency to engaging in Kaldor-Hicks type tradeoffs and assessments of individual utility under various formulations of consumer surplus. These theories can be controversial, but those controversies will not be explored here. For a general discussion of welfare and efficiency analysis, see COOTER & ULEN, *supra* note 25, at 49-51, and POSNER, *supra* note 7, at 12-17. For a discussion of the various approaches to consumer surplus, see *infra* notes 111-120 and accompanying text. For a standard critique of Marshallian consumer surplus, see PAUL A. SAMUELSON, FOUNDATIONS OF ECONOMIC ANALYSIS 195-202 (enlarged ed. 1983).

price competition resulting from merger or cartelization decreases consumer surplus. At the same time, nonprice competition is costly. A reduction in the level of nonprice competition will lead to an increase in producer surplus (profits). If the increase in producer surplus more than offsets the decline in consumer surplus, then the reduction in nonprice competition will yield a net increase in total welfare under the Kaldor-Hicks standard. Thus, an antitrust regime that aims at maximizing total welfare rather than competition or consumer surplus would allow such an exercise of market power.[47]

3. *R&D Joint Ventures, Advertising Restrictions, and Other Examples*

At first blush, intramarket second-best tradeoffs may seem foreign to antitrust analysis. Untamed, second-best analysis certainly has radical potential, and some may appropriately fear opening a Pandora's Box if the presence of such tradeoffs can be used to establish antitrust defenses. In other respects, however, the underlying logic of second-best theory and particular second-best problems are not strangers to antitrust analysis,[48] nor to other fields of law — such as in-

47. Again, the contrast with the consumer surplus standard is telling. If nonprice competition yields any benefit to consumers, then reducing nonprice competition will reduce consumer surplus. A strict consumer surplus standard would forbid any second-best tradeoff in this setting, regardless of the size of the corresponding increase in producer surplus or the net effect on total welfare. There is the possibility that increases in producer surplus could be shared with or "passed on" to consumers. In merger cases, a number of courts require alleged productive efficiency benefits to be passed on in some form to consumers. *See* Hammer, *supra* note 23 (discussing the passing-on requirement and citing cases). Similar analysis would ask if any of the increase in producer surplus associated with reduced nonprice competition would be shared with consumers. Price reductions would be one means. Unfortunately, in markets where inefficient nonprice competition is caused by ineffective price competition, price reductions are an unlikely mechanism to share welfare benefits. *See id.*

48. In the past thirty years, the antitrust standards governing vertical restraints have been completely re-written. The Court has gone from condemning vertical nonprice restraints (i.e., exclusive territories, customer restrictions, and service requirements) as per se illegal to embracing the efficiency-enhancing potential of these agreements under the rule of reason. *See* Continental T.V., Inc. v. GTE Sylvania, Inc., 433 U.S. 36 (1977) (reversing the per se rule condemning vertical nonprice restraints found in *United States v. Arnold Schwinn & Co.*, 388 U.S. 365 (1967), in favor of rule-of-reason evaluation). This transformation resulted from the Court's acceptance of some basic economic theories illustrating the efficiency benefits of vertical restraints. The logic underlying this analysis bears a striking similarity to the logic underlying intramarket second-best tradeoffs. The goal of contemporary antitrust law is the promotion of inter-brand competition. *See* Business Electronics v. Sharp Electronics, Corp., 485 U.S. 717, 726 (1988) ("[I]nterbrand competition is the primary concern of the antitrust laws . . . ."). The Court is willing to accept restrictions on intra-brand competition, i.e., competition between distributors selling the same brand such as Ford automobiles, to the extent that such restrictions increase inter-brand competition, i.e., competition between distributors selling different product brands such as competition between Ford and GM dealers.

The primary justification for vertical nonprice restraints is the prevention of free riding. *See Business Electronics*, 485 U.S. at 731 ("[M]anufacturers are often motivated by a legiti-

tellectual property — which openly accept tradeoffs between the social costs of monopoly and the social benefits of innovation. It is easy to think of numerous additional examples.

a. R&D Joint Ventures. Patents grant state-sanctioned monopolies to inventors for a specified period of time on the theory that monopoly profits are necessary to encourage innovation. The social costs associated with restricting competition and limiting the dissemination of technology are believed to be worth the social benefits of fostering new innovation. This is a type of second-best tradeoff. Technology presents both externalities and public goods problems. It is difficult for the creator of new technology to privately appropriate all the social benefits associated with its creation. Additionally, information is a public good. Consumption by one person does not interfere with consumption of that same information by another person. Social welfare is maximized, ex post, by disseminating information at the near-zero marginal cost of its production. In sanctioning patents and copyrights, the legislature has consciously sacrificed the social value of post-innovation competition for the objective of maximizing total welfare.

These concerns are not limited to the enforcement of intellectual property rights. Antitrust law is frequently called upon to address questions concerning technology. Not surprisingly, these cases can

---

mate desire to have dealers provide services, combined with the reality that price cutting is frequently made possible by 'free riding' on the services provided by other dealers."); *GTE Sylvania*, 433 U.S. at 55 ("Because of market imperfections such as the so-called 'free rider' effect, these services might not be provided by retailers in a purely competitive situation, despite the fact that each retailer's benefit would be greater if all provided the services than if none did."); *see also* Picker International, Inc. v. Leavitt, 865 F. Supp. 951, 968 (D. Mass. 1994) (upholding the reasonableness of vertical restrictions designed to prevent freeriding on substantial R&D investments of a manufacturer of medical equipment).

Free riding is simply one variation of the externality story. A dealer of high-end audio equipment invests substantial resources in providing an elaborate showroom and a trained sales staff. These investments require the dealer to charge a higher price for his equipment to recoup the expenses for nonprice service. The first dealer cannot internalize the benefits of the nonprice investments if a second cost-cutting dealer is permitted to open a shop next door and sell the identical equipment at a lower price because he does not offer the same services. An exclusive territory — an agreement that might otherwise be condemned as a restraint of trade in violation of the Sherman Act — is permitted because it counteracts the undesirable effects of free riding. The exclusive territory results in acknowledged anticompetitive effects. By restricting intra-brand competition, the prices of the branded products are higher than they otherwise would be. Indeed, the higher prices are the intended effect of the restraint, because it permits the provision of services, promotion, and advertising that would not otherwise be possible.

Vertical nonprice restraints are not the same as intramarket second-best tradeoffs. The problem of incomplete contracts underlying vertical restraints is different than the classic market failures subject to intramarket second-best tradeoffs (although many insights can be gained by thinking of contracting failures as a type of market failure). This example reminds us, however, that what we perceive to be foreign and what we perceive to be familiar depends often on framing, and that intramarket second-best tradeoffs can be framed in ways that highlight similarities with established antitrust doctrines.

raise second-best concerns, even if courts do not analyze the issues in terms of second-best theory. Take the example of joint ventures for research and development. Should antitrust courts permit a group of horizontal competitors to form a joint venture aimed at pooling R&D efforts and jointly marketing the resulting products?[49] Multiple mar-

---

49. *See generally* Gene M. Grossman & Carl Shapiro, *Research Joint Ventures: An Antitrust Analysis*, 2 J.L. ECON. & ORG. 315 (1986); Janusz A. Ordover & Robert D. Willig, *Antitrust for High-Technology Industries: Assessing Research Joint Ventures and Mergers*, 28 J.L. & ECON. 311 (1985).

There are surprisingly few published judicial opinions addressing cooperative R&D initiatives. As a general matter, the Supreme Court has indicated that legitimate joint ventures will be favorably treated under the rule of reason. *See* Northwest Wholesale Stationers, Inc. v. Pacific Stationery & Printing Co., 472 U.S. 284, 295 (1985) (suggesting that wholesale purchasing cooperatives are not the type of concerted activity likely to produce anticompetitive effects); NCAA v. Board of Regents, 468 U.S. 85, 113 (1984) (acknowledging the procompetitive potential of joint ventures); Broadcast Music, Inc. v. Columbia Broad. Sys., Inc., 441 U.S. 1, 20-21 (1979) (outlining the procompetitive benefits of BMI's blanket licensing practices). For the most part, oversight of R&D ventures has been the domain of the enforcement agencies and addressed through various enforcement agency guidelines. *See U.S. Dep't of Justice and U.S. Fed. Trade Comm'n Antitrust Guidelines for Collaborations Among Competitors* (Apr. 7, 2000); *U.S. Dep't of Justice and U.S. Fed. Trade Comm'n Antitrust Guidelines for the Licensing of Intellectual Property* (Apr. 6, 1995); and *U.S. Dep't of Justice Antitrust Guidelines for International Operations*, 4 Trade Reg. Rep. (CCH) ¶ 13,109.85 (Nov. 10, 1988) (Illustrative Case 6: Research and Development Joint Venture). Cooperative R&D initiatives are often examined through the business review process. *See, e.g., American Heart Association Pharmaceutical Roundtable*, 1998 DOJBRL LEXIS 7 (Mar. 20, 1998) (initiative sponsoring and funding basic biomedical research); *Amoco, Arco, Exxon, Mobile, Shell, Texaco, Texas A&M University, Initial Members of Cooperative*, 1997 DOJBRL LEXIS 9 (Apr. 23, 1997) (cooperative project designed to perform exploration and production research of a nature that is not attractive to individual firm research).

The handful of judicial opinions that have examined cooperative R&D efforts fail to provide a systematic or comprehensive approach to the problem. The classic example of collusive under-investment in R&D is the 1960s auto industry conspiracy to suppress research in pollution control devices. *See* Washington v. General Motors Corp., 406 U.S. 109, 111-12 (1972) (outlining the elements of the antitrust conspiracy). Ironically, one of the alleged justifications for the conspiracy was fear on the part of the individual auto companies that they would be unable to recoup the cost of their investment in private pollution control due to the externalities associated with environmental problems. *See In re* Multidistrict Vehicle Air Pollution M.D.L. No. 31, 481 F.2d 122 124 (9th Cir. 1973) ("Appellees urge that this horizontal antitrust conspiracy was motivated ... by appellant's conviction that antipollution devices are externalities, whose development would increase price without concomitant spur to consumer interest."). Other courts have expressed doubts as to whether alleged collusive efforts of tobacco companies to suppress research and development or the marketing of new products would constitute an antitrust violation. *See* Steamfitters Local 420 v. Philip Morris, Inc., 171 F.3d 912, 925 n.7 (3d Cir. 1999) ("[I]t is not clear whether even a concerted decision among all of the businesses in an industry to keep one of their new products from reaching consumers would be an antitrust violation."). Conflicting judicial statements can also be found concerning the desirability of competition in the R&D arena. *See, e.g.*, SCFC ILC, Inc. v. Visa U.S.A., Inc., 819 F. Supp. 956, 981-83 (D. Utah 1993) (noting the range of legislative and policy arguments that might favor cooperative R&D initiatives); Universal Analytics v. MacNeal-Schwendler Corp., 707 F. Supp. 1170, 1179 (C.D. Cal. 1989) (expressing reluctance to treat certain R&D activities as predatory for antitrust purposes for fear of chilling future innovation); United States v. Western Elec. Co., 673 F. Supp. 525, 559-61 (D.D.C. 1987) (concluding that competition between the Regional Bell Operating Companies has improved research, development, innovation, and the introduction of new technology); North Carolina v. Chas Pfizer & Co., 384 F. Supp. 265, 284 n.28 (E.D.N.C. 1974) (questioning the desirability of competition in terms of R&D in the pharmaceutical indus-

ket failures are present. The public good aspects of information suggest that there is likely to be a suboptimal amount of industry R&D. The public good characteristics also suggest that a direct government subsidy may be a legitimate policy substitute for permitting increased private cooperation. R&D also raises free rider problems. The research results may not be patentable, or may be only imperfectly protected by patents and other intellectual property rights. No single firm may want to invest substantial funds in R&D for fear that its competitors may simply copy the results and market a product at a lower price, given that they did not have to absorb the initial R&D costs.

A joint venture is one means of internalizing some of the externalities and could result in higher levels of R&D than would otherwise exist. At the same time, these joint ventures can raise substantial antitrust concerns. Even if we can be certain that there will be a higher aggregate level of R&D because of the joint venture,[50] the venture is likely to produce market power in the newly created product market that would not have existed under a regime of competitive R&D. Competition, however, might lower the probability of any innovation taking place at all, or might substantially delay the innovation. Whether competition or cooperation will increase total welfare is often a difficult question,[51] but a question that may be better answered if

try). Similar issues have been raised by administrative bodies charged with assessing the competitive effects of the conduct of various regulated entities. *See, e.g.*, Money Stations v. Board of Governors of Fed. Res. Sys., 81 F.3d 1128, 1134-36 (D.C. Cir. 1996) (Board's determination under the Bank Holding Act that the consolidation of ATM networks would produce public benefits in terms of expanded R&D efforts was not supported by substantial evidence in the record); Ft. Pierce Utilities Authority v. United States, 606 F.2d 986, 990 (D.C. Cir. 1978) (discussing early legislative exemption from regulatory antitrust scrutiny for nuclear power plants as "research and development" facilities).

50. Cooperative R&D can raise anticompetitive risks. Under some circumstances, competition may generate stronger R&D incentives than will cooperation. There is also a risk of collusive under-investment in R&D. *See* William F. Baxter, *The Definition and Measurement of Market Power in Industries Characterized by Rapidly Developing and Changing Technologies*, 53 ANTITRUST L.J. 717, 720 (1984); Grossman & Shapiro, *supra* note 49, at 324. Additional anticompetitive risks of cooperative R&D are explored *infra* note 94 and accompanying text.

51. The effect that cooperation and competition in R&D can have on total welfare has been examined by economists. *See, e.g.*, Barry Bozeman et al., *An Economic Analysis of R&D Joint Ventures*, 7 MANAGERIAL & DECISION ECON. 263 (1986) (predicting that cooperation will increase aggregate amounts of R&D in a model emphasizing the public good nature of basic research); Alexis Jacquemin, *Cooperative Agreements in R&D and European Antitrust Policy*, 32 EUR. ECON. REV. 551, 553-57 (1988) (examining the potential public benefits and possible public costs of cooperative R&D initiatives); Michael L. Katz, *An Analysis of Cooperative Research and Development*, 17 RAND J. ECON. 527, 536-38 (theoretical assessment of the welfare effects of cooperative R&D); Takeo Nakao, *Cost-Reducing R&D in Oligopoly*, 12 J. ECON. BEHAV. & ORG. 131, 145-46 (1989) (highlighting the potentially ambiguous welfare effects of competitive and cooperative R&D efforts).

the second-best dimensions underlying the problem are expressly acknowledged.[52]

b. Advertising Restrictions. Information market failures also raise troubling questions for antitrust courts. Information in markets may be imperfect because some characteristics of the product may be unknown to consumers, or can be acquired by consumers only after engaging in costly searches. It may be incomplete because purchasers may not know all the price and quality attributes of competing products. And information may be asymmetrically held because sellers may have substantially better information than buyers. Information problems can confront producers as well as consumers. Competitors can respond to information problems in different ways. They can share known information or they can cooperate in generating new information. Competitors can engage in self-regulation to prevent agency problems resulting from asymmetric information. Alternatively, competitors can attempt to control the type of advertising that takes place in a world of imperfect information. Whether courts expressly think of these issues in terms of second-best tradeoffs or not, these cases are decided in the shadow of market failures and second-best concerns.

The Supreme Court's recent decision in *California Dental Ass'n v. F.T.C.* is illustrative.[53] At issue was a practice of the nonprofit California Dental Association restricting "price advertising, particularly discounted fees, and advertising related to the quality of dental services."[54] The Federal Trade Commission found the practice to be a per

52. How an affirmative defense for intramarket second-best tradeoffs would treat R&D joint ventures is examined *infra* notes 93-94 and accompanying text.

53. California Dental Ass'n v. FTC, 119 S. Ct. 1604 (1999). The articles relied upon by the Court provide a good introduction to the problem of informational market failures, particularly in markets for professional services. *See id.* at 1613-14 (discussing information problems in the market for dental services and citing Jack L. Carr & Frank Mathewson, *The Economics of Law Firms: A Study of the Legal Organization of the Firm*, 33 J.L. & ECON. 307 (1990); George A. Akerlof, *The Market for "Lemons": Quality Uncertainty and the Market Mechanism*, 84 Q.J. ECON. 488 (1970); Hayne E. Leland, *Quacks, Lemons, and Licensing: A Theory of Minimum Quality Standards*, 87 J. POL. ECON. 1328 (1979); 1 B. FURROW, T. GREANEY, S. JOHNSON, T. JOST & R. SCHWARZ, HEALTH LAW § 3-1, p. 86 (1995); and Robert G. Evans, *Professionals and the Production Function: Can Competition Policy Improve Efficiency in the Licensed Professions?*, *in* OCCUPATIONAL LICENSURE AND REGULATION 225 (Simon Rottenberg ed., 1980)). For a discussion of the impact of imperfect information on the functioning of competitive markets, see generally Alan Schwartz & Louis Wilde, *Intervening in Markets on the Basis of Imperfect Information: A Legal and Economic Analysis*, 127 U. PA. L. REV. 630, 641-51 (1979) (summarizing the results of economic models that examine the behavior of markets characterized by imperfect information); Richard Craswell, *Interpreting Deceptive Advertising*, 65 B.U. L. REV. 657, 686-88 (1985) (examining the effects of deceptive advertising on competitive equilibria); and Lillian R. BeVier, *Competitor Suits for False Advertising Under Section 43(a) of the Lanham Act: A Puzzle in the Law of Deception*, 78 VA. L. REV. 1, 4-13 (1992) (reviewing economic literature on the competitive effects of advertising).

54. 119 S. Ct. at 1609.

se violation of antitrust law, and in the alternative to be an unreasonable restraint under a quick-look rule-of-reason analysis. The Ninth Circuit affirmed on the basis of abbreviated rule-of-reason analysis. The Supreme Court reversed. Agreements, it said, are presumed to be unlawful under the quick-look rule of reason if "an observer with even a rudimentary understanding of economics could conclude that the arrangements in question would have an anticompetitive effect on consumers and markets."[55] This question, according to the Court, must be answered in light of the "striking disparities between the information available to the professional and the patient."[56]

The Court proceeded to catalogue the informational failures plaguing the market for dental services. Information in markets for professional services is asymmetrically held — dentists are better informed than their patients. Moreover, it is difficult for patients to get and verify information because of the search costs involved and the experiential nature of some services. With experience goods such as dental services, consumers are unable to assess the quality of the dentist's skills until after they agree to and receive the service. Certain dental services are also credence goods, meaning that even after receiving the service, consumers may not be able to accurately assess the quality of the dentist:

> [T]he quality of professional services tends to resist either calibration or monitoring by individual patients or clients, partly because of the specialized knowledge required to evaluate services, and partly because of the difficulty in determining whether, and the degree to which, an outcome is attributable to the quality of the service (like a poor job of tooth filling) or to something else (like a very tough walnut).[57]

In the presence of multiple market failures, conduct that might otherwise appear to be anticompetitive may in fact be welfare-enhancing. The California Dental Association argued that the advertising restrictions were procompetitive because they limited false and misleading advertising, even if they also had the unintended effect of limiting some truthful or desirable information as well. The Court found the procompetitive potential of the restraint sufficiently plausible to withstand quick-look rule-of-reason analysis and to require a full-scale rule-of-reason examination:

> Put another way, the CDA's rule appears to reflect the prediction that any costs to competition associated with the elimination of across-the-board advertising will be outweighed by gains to consumer information (and hence competition) created by discount advertising that is exact, accurate and more easily verifiable (at least by regulators). As a matter of

55. 119 S. Ct. at 1612.

56. 119 S. Ct. at 1613.

57. 119 S. Ct. at 1613-14.

> economics this may or may not be correct, but it is not implausible, and neither a court nor the Commission may initially dismiss it as preemptively wrong.[58]

It remains an open question how best to evaluate the desirability of these types of restrictions under the rule of reason.[59]

c. Other Illustrations. R&D joint ventures and restrictions on advertising involve cases where antitrust courts and the enforcement agencies are already struggling with market failures in the context of rule-of-reason analysis. The following examples illustrate other areas where intramarket second-best tradeoffs might be found. These examples suggest more novel claims. Philip Areeda provides the hypothetical (a little dated in the age of cable TV and VCRs) of the major television networks agreeing that each will set aside one hour an evening for "quality" programming — cultural programming aimed at audiences too small to be supported by commercial advertising.[60] Could such a collusive restraint withstand antitrust scrutiny? The agreement might be reasonable if it could be characterized as counteracting market failures:

> The economic argument would be that the television industry exhibits this "market failure": given that government allocation limits the number of channels and that each station or network seeks to maximize its advertising revenues and therefore its audience for each broadcast hour, cultural services that viewers and advertisers are ready to support are not offered (unlike newspapers which can simultaneously provide culture and comic strips). It would then be argued that the time set-aside for quality corrects this market failure and thereby brings about a more "competitive" result.[61]

This example arises in the context of a regulated industry. One can imagine additional second-best concerns emerging in those sectors of the economy that are currently being deregulated — telecommunications, natural gas, and electricity. Relying upon competition in settings that retain quasi-public utility characteristics may engender forms of cooperation that facially appear to be anticompetitive, but which may have underlying efficiency justifications.

Other areas of second-best tradeoffs may involve the relationship between market structure and the appropriate tort liability standard,[62]

---

58. 119 S. Ct. at 1615.

59. How an affirmative defense for intramarket second-best tradeoffs would treat advertising restrictions is examined *infra* notes 95-97 and accompanying text.

60. *See* PHILIP E. AREEDA, THE "RULE OF REASON" IN ANTITRUST ANALYSIS: GENERAL ISSUES 7-8 (Federal Judicial Ctr., Educ. and Training Series, 1981).

61. *Id.* at 8.

62. *See* A. Mitchell Polinsky & William P. Rogerson, *Product Liability, Consumer Misperceptions, and Market Power*, 14 BELL J. ECON. 581 (1983) (exploring the relationship between market structure and the appropriate standard of tort liability).

or the relationship between the number of competitors and price in markets where search costs and reputation play a significant role,[63] or the entry into a monopoly market by a higher cost producer.[64] Finally, one can imagine numerous types of rent-dissipating behavior that could justify second-best analysis. Industries may engage in advertising which is excessive by the total welfare criterion because it is persuasive without being informative. Pharmaceutical companies, for example, may devote too much money to influence the prescription practices of physicians, or markets may exhibit too much product differentiation.[65]

---

63. *See* Mark V. Pauly & Mark A. Satterwaite, *The Pricing of Primary Care Physicians' Services: A Test of the Role of Consumer Information*, 12 BELL J. ECON. 488, 489 ("First, if the number of physicians within a community increases, then consumer information about each physician decreases; thus, consumers have a more difficult time in the search for new physicians. Second, if this search becomes more difficult, then consumers become less price sensitive, i.e., each physician's demand curve becomes less elastic. Consequently, an increased supply of physicians, or an increase in any other factor that makes consumers' search more difficult, may cause physician equilibrium fees to rise.").

64. Richard Schmalensee has written about this possibility. *See* Richard Schmalensee, *Is More Competition Necessarily Good?*, 4 INDUS. ORG. REV. 120 (1976). Where the phenomenon Schmalensee identifies exists, entry into a monopoly market by a higher-cost competitor may actually reduce total welfare. Two considerations are at play. So long as the producers act independently, entry will result in an increase in output and a decrease in price, which will unambiguously increase consumer surplus. The higher-cost producer, however, is less efficient. At some point, depending upon how much higher the entrant's costs are and the amount of output being produced by the higher-cost firm, the social loss associated with the higher costs and the reduction in producer surplus can outweigh the increase in consumer surplus, leading to a reduction in total welfare. If this is the case, then entry may not be desirable, and government restrictions on entry, or the sanctioning of conduct that may otherwise be predatory or exclusionary may be justified. *See id.* at 120 ("The apparent policy implication of this result is disturbing: privately profitable entry may not be socially desirable if the entrant's costs exceed those of existing firms. Society as a whole would be better off if existing firms would be allowed to bribe potential entrants not to enter, or if entry was restricted by government regulation of some sort."). This example again illustrates the difference between a total welfare and a consumer surplus standard. Under a total welfare standard, entry would be prohibited if the loss in producer surplus exceeded the gain in consumer surplus. Under a consumer surplus standard, entry would be permitted if there would be any increase in consumer surplus.

65. It is appropriate to end this list of examples with one that sounds a cautionary note. Second-best analysis invites the pursuit of interesting theoretical puzzles. It is one thing to present a plausible second-best story; it is quite another to substantiate the theory and to present evidence of a real increase in social welfare. Insurance markets, for example, are associated with problems of moral hazard. Moral hazard is typically associated with higher than optimal rates of consumption. It is plausible that market power (by increasing price and decreasing consumption) may counteract this tendency and lead to a net increase in total welfare. Hence, if second-best analysis had a place at the antitrust table, merging providers in such industries might have a defense to charges of monopolizing the market.

Yet, in health care markets where this theory has been examined, it appears that this intuition is incorrect. Martin Gaynor, Deborah Haas-Wilson, and William Vogt demonstrate that increasing the price of medical services through increased provider market power does not necessarily increase total welfare (counteract the effects of moral hazard) in the face of a competitive insurance market. *See* Martin Gaynor et al., Are Invisible Hands Good Hands? Moral Hazard, Competition, and the Second Best in Health Care Markets, at 3 (NBER Working Paper No. W6865, Dec. 1998) (on file with author). The intuition behind this analysis is relatively straightforward. Insurance companies are already acting to establish

How should courts respond to these claims that market power may be desirable? At this juncture, it is useful to distinguish two distinct claims advanced in this Article. My primary claim is that the antitrust framework should be broadened to consider these types of intramarket second-best tradeoffs. Private conduct that ordinarily would be condemned as a restraint of trade, either as single-firm behavior or as collaborative behavior, may in fact be desirable. What makes the prima facie bad behavior good is that it occurs in a market that is imperfect, that in one or more ways has "failed" and is not likely to correct its failures on its own. Despite the seeming paradox, the undesirable consequences of market failure may be ameliorated by private behavior that in other circumstances would itself be a market failure. My second claim is that courts should assess the desirability of such tradeoffs in reference to the effect of the restraint on total welfare. This requires examining the well-being of producers along with the well-being of consumers. While I believe the claims are complementary, it is possible to acknowledge the relevance of intramarket second-best concerns and assess such tradeoffs in light of their effect on consumer surplus.

The argument that under certain circumstances market power may produce beneficial results, and that antitrust law should be flexible enough to accommodate such possibilities is not entirely new. The next Section examines how intramarket second-best concerns compare with the productive efficiency defense advocated by Oliver Williamson, and explores the similarities between second-best tradeoffs and market failure defenses that have been proposed by various antitrust scholars.

---

optimum rates of co-insurance and consumer cost-sharing in light of the prevailing moral hazard concerns. *See id.* at 3. Optimally set co-insurance rates will not necessarily yield anything close to the first-best outcome, but they do imply that further welfare gains cannot be attained along this dimension by further raising prices through increased provider market power.

It is significant that Gaynor, Haas-Wilson, and Vogt examined the market failure of moral hazard in isolation. As the discussion of hospital non-price competition in Section I.B.2 suggests, increased provider market power may result in welfare gains unrelated to the problem of moral hazard. Moreover, different market failures, such as those involving non-price competition, or different combinations of market failures may well produce welfare effects different from the ones identified by these authors. *See id.* at 20 ("We must apply some caveats to these conclusions, however. In this paper, we have analyzed one of the distortions in medical markets: moral hazard. We have not considered other factors that are commonly cited in rendering competition in medical markets different: risk selection in insurance markets, agency problems in medical markets (i.e., induced demand), and the presence of non-profit firms. It remains for future research to consider the constellation of these imperfections in concert.").

C. *Antitrust Law and Second-Best Concerns: From Productive Efficiency to Market Failure*

1. *Williamson's Productive Efficiency Defense*

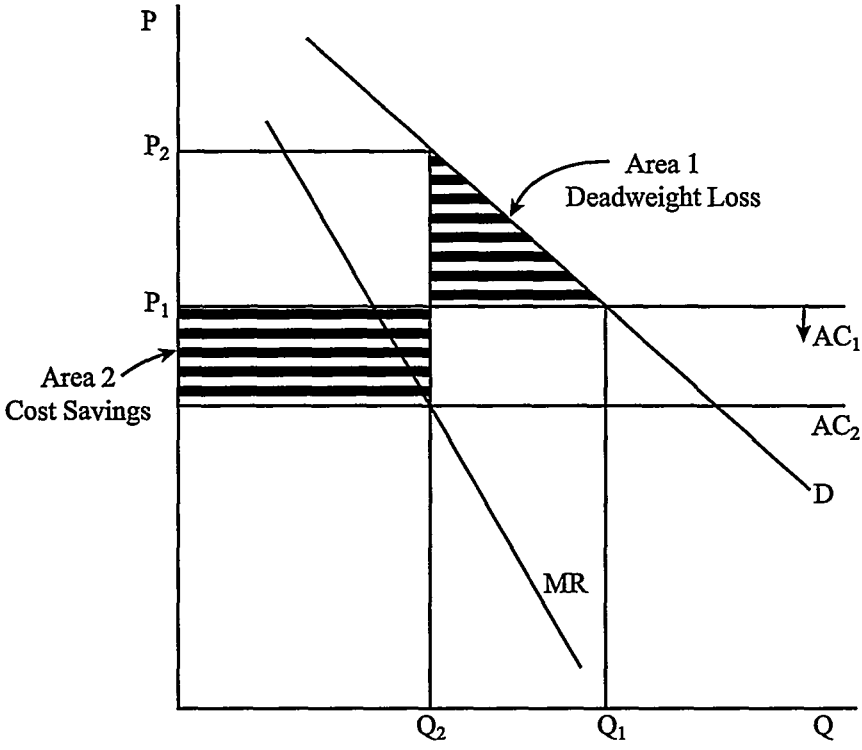
Analytically, the issues raised by intramarket second-best tradeoffs are similar to those raised by the efficiency defense long advocated by Oliver Williamson.[66] Williamson recognized that certain forms of productive efficiency can be achieved only at the price of sacrificing allocative efficiency.[67] Figure 3 reproduces the partial equilibrium framework Williamson developed to assess these tradeoffs.[68]

66. *See generally* Oliver E. Williamson, *Economies as an Antitrust Defense: The Welfare Tradeoffs*, 58 AM. ECON. REV. 18, 18 (1968) [hereinafter Williamson, *Economies as an Antitrust Defense*]; Williamson, *supra* note 2, at 105; Oliver E. Williamson, *Economies as an Antitrust Defense: Correction and Reply*, 58 AM. ECON. REV. 1372 (1968); Oliver E. Williamson, *Economies as an Antitrust Defense: Reply*, 59 AM. ECON. REV. 954 (1969); Oliver E. Williamson, *Economies as an Antitrust Defense Revisited*, 125 U. PA. L. REV. 699, 699 (1977) [hereinafter Williamson, *Economies Revisited*].

67. "A production process is said to be *productively efficient* if it yields a given level of output with the least cost combination of inputs." COOTER & ULEN, *supra* note 25, at 17. "A particular distribution of goods among consumers is said to be *allocatively efficient* if it is not possible to redistribute the goods so as to make at least one consumer better off (in his own estimation) without making another consumer worse off (again, in his own estimation)." *Id.* at 18.

68. For Williamson's exposition of the model and the efficiency tradeoffs, see Williamson, *Economies Revisited*, *supra* note 66, at 706-09, and Williamson, *Economies as an Antitrust Defense*, *supra* note 66, at 21-23. For summaries or discussions of Williamson's model, see BORK, *supra* note 12, at 107-10; Jerome Ellig, *Computer Reservation Systems, Creative Destruction, and Consumer Welfare: Some Unsettled Issues*, 19 TRANSP. L.J. 287, 295-96 (1991); Fisher & Lande, *supra* note 23, at 1626-30; and David Millon, *The Sherman Act and the Balance of Power*, 61 S. CAL. L. REV. 1219, 1229-30 n.43 (1988). Williamson's views have evolved since the 1960s to stress the importance of institutional economics and transaction costs. *See, e.g.*, OLIVER E. WILLIAMSON, MARKETS AND HIERARCHIES: ANALYSIS AND ANTITRUST IMPLICATIONS (1975); OLIVER E. WILLIAMSON, THE ECONOMIC INSTITUTIONS OF CAPITALISM (1985). Still, Williamson's partial equilibrium framework provides an appropriate means of examining the static efficiency of mergers. Comparative institutional analysis can be folded into the dynamic efficiency considerations outlined in the affirmative defense discussed *infra* Section I.D.

FIGURE THREE: WILLIAMSON'S EFFICIENCY TRADEOFFS



Assuming that the increase in productive efficiency could be obtained only by a substantial increase in economic concentration (a merger to monopoly in the extreme case), the cost savings association with increased productive efficiency (represented graphically by a downward shift in the producer's cost function — "Area 2 Cost Savings" in figure 3) is balanced against the loss in allocative efficiency represented by the deadweight loss triangle traditionally associated with monopolies ("Area 1" in figure 3). If the gains in productive efficiency exceed the losses from decreased allocative efficiency, then, Williamson argued, antitrust law should permit the increase in market power.[69]

69. Williamson combined this framework for analyzing efficiency tradeoffs with the predictive insight that relatively small cost reductions would outweigh losses in allocative efficiency. *See* Williamson, *Economies Revisited*, *supra* note 66, at 699, 709; Williamson, *Economies as an Antitrust Defense*, *supra* note 66, at 22-23. Productive efficiency gains occur with every product produced while the loss of allocative efficiency is limited to marginal consumers. This same logic underlies the assertion that the "wealth transfer" effects of market power (the transformation of consumer surplus into producer surplus) will also normally outweigh the deadweight loss. *See* Fisher & Lande, *supra* note 23, at 1644-46; Harris & Jorde, *supra* note 13, at 11-12.

The form of Williamson's argument is similar to the form of the argument in favor of intramarket second-best tradeoffs. Conduct that would otherwise be viewed as anticompetitive and therefore prohibited by the antitrust laws (the accumulation of market power) should be permitted because the conduct produces an increase in total welfare. In addition, both the analysis of productive efficiency claims and intramarket second-best tradeoffs adopt a partial equilibrium framework and employ a total welfare standard to guide antitrust decisionmaking. This last point deserves emphasis. Commentators often overlook the fact that Williamson's graphic illustration of the tradeoff between allocative and productive efficiency is simply a means of ascertaining the effect of merger on total welfare. In his 1969 and 1977 treatments of efficiency tradeoffs, Williamson introduced a basic welfare function to complement the graphic analysis:

> The partial equilibrium apparatus that I would propose for purposes of examining the trade-off question is one in which the welfare function is expressed as $W = (TR + S) - (TC - R)$, where TR refers to total revenue, S to consumer surplus, TC to total cost, and R to intramarginal rents. The terms in the first set of parentheses reflect the social benefits associated with the activity in question, while the terms in the second (under appropriate restrictions) reflect social costs. The allocative efficiency consequences of any merger that increases both efficiency and market power can be evaluated only by estimating net effects.[70]

The implications of the alleged restraint on allocative and productive efficiencies are simply subcomponents in a broader welfare analysis. The ultimate focus is the impact of a proposed merger or other restraint of trade on total welfare.[71]

Williamson's analysis, however, presents less difficult doctrinal issues for antitrust scholars than do intramarket second-best tradeoffs. While Williamson's total welfare orientation is broad enough to encompass a range of other allocative efficiency concerns, he focuses on the effects of market power on productive efficiency. This focus has its antecedents in concerns over economies of scale, natural monopolies, and the problems presented by the potentially efficient monopolist. The proposition that antitrust law should not be interpreted in a

70. Williamson, *supra* note 2, at 107; *see also* Williamson, *Economies Revisited*, *supra* note 66, at 708 n.27.

71. Williamson's framework is also useful because the arguments that have been levied for and against his efficiency tradeoffs parallel many of the arguments that can be made for and against permitting intramarket second-best tradeoff, or total welfare analysis more generally. *See, e.g.*, Fisher & Lande, *supra* note 23, at 1624-51 (rejecting Williamson's proposed tradeoff analysis and its underlying total welfare standard because of the assorted difficulties associated with litigating the defense).

manner that needlessly sacrifices productive efficiency has a historically respectable pedigree.[72]

Intramarket second-best analysis, however, is concerned not with the implications of market power for productive efficiency, but with the effects of market structure on allocative efficiency. Analytically, second-best concerns are the flip side of the same coin as Williamson's productive efficiency defense, and both fit comfortably within a total welfare orientation. Pragmatically, however, a market failure defense is more controversial. Intramarket second-best tradeoffs more profoundly question the assumption that "competition" is either workable or desirable. Hence, the argument that intramarket second-best tradeoffs can justify restraints of trade has substantially less support in the history and tradition of antitrust enforcement than do productive efficiency claims.

2. *Market Failure Defenses*

The possibility of a market failure defense has been waiting in the wings of antitrust law for many years. Phillip Areeda posed the question: "[C]an some general welfare claims be expressed as improvements in competition?"[73] Areeda thought they could be, if they corrected market failures and were associated with an increase in total welfare, such as an expansion of output.[74] "We might even believe that the apparent restraint actually moves market performance closer to the competitive result. Rather than suppressing competition, offsetting a 'market failure' promotes competitive results."[75] Doctrinally, Areeda recasts remedying market failures and increasing total welfare in the rhetoric of "promoting competition."[76]

---

72. A number of these issues were recognized and debated during the passage of the Sherman Act, and these same issues have occupied American economists since the 1880s. *See generally* F.M. Scherer, *Efficiency, Fairness and the Early Contribution of Economists to the Antitrust Debate*, 29 WASHBURN L.J. 243 (1990). To be sure, at the time of Williamson's writings, many of these lessons had been lost on a Court that seemed to treat certain productive efficiencies as a reason for opposing mergers. *See, e.g.*, FTC v. Procter & Gamble Co., 386 U.S. 568, 580 (1967) ("Possible economies cannot be used as a defense to illegality."); Brown Shoe Co. v. United States, 370 U.S. 294, 344 (1962) ("Congress appreciated that occasional higher costs and prices might result from the maintenance of fragmented industries and markets. It resolved these competing considerations in favor of deconcentration."). In this respect, Williamson's work was an important wake-up call. Today, most contemporary approaches to antitrust law would acknowledge a need to accommodate productive efficiency in some fashion, even if they disagree over the specific form of the accommodation.

73. PHILLIP E. AREEDA, ANTITRUST LAW § 1504, at 382 (1978); *see also* AREEDA, *supra* note 60, at 7-8 (exploring the possibility of justifying restraints of trade that are targeted at correcting market failures as being procompetitive).

74. *See* AREEDA, *supra* note 73, § 1504, at 383.

75. *Id.*

76. *Id.* ("This is not the place to work out the soundness of that argument. It is enough to see that many claims of redeeming virtue expressed by laymen in 'public interest' terms can be reformulated in terms of promoting competition."); *see also id.* § 1511, at 434-36 (ex-

Thomas Greaney has built on Areeda's insight to advocate a market failure defense in health care.[77] Greaney focuses on the "procompetitive" objective of overcoming the market failures plaguing the health care industry. He contends that "the defense could signify nothing more than an application of ancillary restraints principles."[78] Recognizing that defendants could manipulate a market failure defense, he proposes a number of limiting principles. First, his defense would be narrowly construed if the underlying conduct constituted a "naked restraint."[79] Second, "courts should require proof that removal of the market imperfections would actually improve competitive outcomes."[80] Finally, courts must "assess the risk that self-regulatory power might be abused or would contribute to anticompetitive conditions."[81] Into this last principle Greaney reads a requirement that the restraint be necessary to the procompetitive objective and be no broader in scope than required to achieve the objective.[82]

Joseph Brodley similarly advocates allowing a defense for "productive" and "innovative" efficiencies, but defines these efficiencies in terms of overcoming competitive market failures.[83] "[I]n cases of competitive market failure, antitrust should recognize a limited efficiencies justification for transactions that temporarily restrict competition in order to correct market failure and thereby advance production and innovation efficiency."[84] Brodley would impose three conditions on such a defense:

---

amining the increasing range of factors considered by the Court to constitute "procompetitive" justifications under the rule of reason); PHILLIP E. AREEDA & HERBERT HOVENKAMP, ANTITRUST LAW § 1504, at 397-402 (Supp. 1998) (exploring the growing range of "legitimate objectives" recognized by lower courts engaging in rule-of-reason analysis).

77. *See* Thomas L. Greaney, *Quality of Care and Market Failure Defenses in Antitrust Health Care Litigation*, 21 CONN. L. REV. 605, 627-49 (1989). Market failures in health care include informational problems, externalities, the presence of public goods, and the effects of insurance. *See id.* at 632-40. Other scholars have acknowledged the possibility of a market failure defense to justify forms of professional self-regulation, but have cautioned against too readily recognizing such a defense in health care markets. *See* Clark C. Havighurst, *Doctors and Hospitals: An Antitrust Perspective on Traditional Relationships*, 1984 DUKE L.J. 1071, 1095-96; Clark C. Havighurst, *Professional Peer Review and the Antitrust Laws*, 36 CASE W. RES. L. REV. 1117, 1143-44 (1986); Clark C. Havighurst & Nancy M.P. King, *Private Credentialing of Health Care Personnel: An Antitrust Perspective*, 9 AM. J.L. & MED. 263, 297-98 (1983).

78. Greaney, *supra* note 77, at 629.

79. *See id.* at 647.

80. *Id.*

81. *Id.* at 648.

82. *See id.*

83. *See* Brodley, *supra* note 21, at 1020-21.

84. *Id.* at 1042. "The necessary precondition for an efficiencies defense is competitive market failure — a condition in which interfirm rivalry does not enhance social wealth and advance the long-run interest of consumers." *Id.* at 1046. "A competitive market failure

> First, the projected activity must increase total social welfare by realizing significant production or innovation efficiencies. Second, the activity must be necessary to achieve such efficiencies and, among reasonably available alternatives, be least harmful in its effect on consumers. Third, the activity must not permanently suppress interfirm rivalry, but must allow for the eventual reestablishment of competition.[85]

Brodley finds a doctrinal home that is similar to the "procompetitive" rationalizations of Areeda and Greaney and notes that in recent years, the Court "has been more willing to uphold private action involving the creation or exercise of market power as an appropriate response to market failure."[86] "In the presence of either type of market failure the recognition of an efficiencies defense creates no tension between promoting socially desirable competition and achieving efficiency because, in the absence of the efficiency-justified restraint, there would be less competition still."[87] Remedying market failures, increasing efficiency, and improving social welfare are all, in Brodley's view, "procompetitive."

A final advocate of a market failure defense is Christopher Leslie.[88] Leslie proposes a market failure defense to horizontal price fixing claims where the commodity at issue is associated with substantial negative externalities. "To warrant a successful market failure defense, a product must possess four characteristics: (1) it must have negative externalities, (2) it must be undertaxed, (3) it must have elastic demand, and (4) it must have feasible alternatives."[89] This approach focuses on the particular economic characteristics of the product, rather than on the welfare effects of the restraints, although the factors identified by Leslie would each be relevant in assessing the magnitude of the pre-restraint loss in allocative efficiency and the post-restraint increase (if any) in total welfare. Leslie's primary contribution, however, is to recognize the serious doctrinal challenges that face any effort to establish a market failure defense and to appreciate

---

requires a breakdown in one or more of the basic conditions for a competitive market . . . ." *Id.* Brodley gives examples of externalities, *see id.* at 1046-47, and of the "inability of an inventor to capture economic reward in the absence of a monopoly patent right," *id.* at 1047.

85. *Id.* at 1037-38. Procedurally, Brodley's defense would entail a two-stage administrative process — a threshold or ex ante stage followed by an ex post review. *See id.* at 1048. At the ex ante stage, the defendants must show that the alleged efficiencies are plausible, that the conduct represents the least restrictive means of addressing the market failure, and that the court could prospectively impose an effective remedy if the conduct failed to achieve the projected efficiencies. *See id.* At the ex post stage, the court would assess whether the efficiencies had in fact been obtained. *See id.* at 1048-49.

86. *Id.* at 1047 (citing Broadcast Music, Inc. v. Columbia Broad. Sys., Inc., 441 U.S. 1 (1979)).

87. *Id.*

88. *See* Leslie, *supra* note 24, at 243.

89. *Id.* at 275.

the fact that there may be a substantial difference between conduct that is "efficient" in the sense of being welfare-enhancing and conduct that is "competitive" or "procompetitive."[90]

D. *The Anatomy of an Affirmative Defense for Intramarket Second-Best Tradeoffs*

My case for intramarket second-best claims follows in the tradition of these writers. I believe that claims for second-best tradeoffs should be treated as an affirmative defense capable of rebutting the legal presumption that conduct is anticompetitive. One setting where such a defense might matter is where a merger between firms produces post-merger concentration ratios that exceed enforcement agency and court-established thresholds, but where such market power may actually be welfare-enhancing. Another setting is where conduct traditionally considered to be per se illegal — price fixing and horizontal territorial divisions — and therefore presumed to be anticompetitive, may in fact have beneficial consequences.

On the theory that the core concern of the antitrust laws is overall market efficiency, this antitrust defense focuses more on total welfare than competition. To make out a successful intramarket second-best defense, defendants in my view would have to establish the following: first, that the alleged anticompetitive conduct remedies an identifiable market failure or market failures; second, that the conduct will result in a net increase in total welfare (static efficiency);[91] third, that the conduct will not substantially impair the ability of public or private actors subsequently to ameliorate the effects of the market failure (dynamic efficiency); and finally, that there is not a less restrictive alternative consistent with the antitrust laws that the parties could undertake to achieve similar welfare gains.[92] In the initial stages and until further

90. *See id.* at 255-67 (examining the Supreme Court's treatment of efficiency and competition). In addition, Leslie provides a well-developed discussion of the justifications for a market failure defense within the context of the second-best theory. *See id.* at 267-72. Leslie's characterization of the conflict between competition and efficiency differs substantially from the harmonious views of Areeda, Greaney, and Brodley. In important respects, however, both views are correct. Recognizing a market failure defense is at the same time a natural evolution of some approaches to antitrust law *and* a radical substitution of economic efficiency in place of competition for other schools of antitrust thought. This underscores the doctrinally provocative side of intramarket second-best concerns and anticipates the discussion of antitrust scholarship and methodology that is examined *infra* Part III.

91. If one wanted to address intramarket second-best concerns but retain an antitrust focus on consumer surplus, this factor could be redefined in terms of proving an increase in consumer surplus.

92. There are important similarities and differences between the defense I propose and the market failure defenses examined in the previous Section. Each defense is rooted in the existence of economic market failures and private efforts that arguably overcome those failures. No level of altruism, however, should be attributed to the private actors. Private parties act to maximize profits. Intramarket second-best tradeoffs and other market failure de-

evidence exists as to the ability to litigate and manage intramarket second-best claims, defendants should be required to demonstrate the four elements with clear and convincing evidence. These rigorous standards recognize that allowing the sort of second-best defense that I advocate liberalizes the status quo and cautions that until the merits of the defense are demonstrated in actual cases, errors should be those that disproportionately stem from doing what we do now.

Some might argue that the test is too rigorous and that the defense will forever remain hypothetical, as no merger or restraint of trade will be able to meet my proposed standards. This is not true. The cases with second-best overtones, such as R&D joint ventures, that currently pass antitrust muster under the rule of reason illustrate that strong second-best claims are not merely hypothetical. The public good and positive externalities characteristics of R&D joint ventures are plain. It should be relatively easy in appropriate cases to demonstrate an increase in total welfare (static efficiency) by documenting increased levels of research expenditures, increased probabilities of successful innovation, and the likelihood that new products will be

---

fenses seek to exploit the potential coincidence of actions that increase both producer and total welfare, while still enforcing antitrust prohibitions in those cases where private actions decrease total welfare.

Although each defense seeks to identify private conduct that is efficiency-enhancing, the intramarket second-best defense, unlike those of Areeda and Greaney, accomplishes this by requiring parties to prove a net increase in total welfare, rather than showing that the conduct "promotes competitive results," AREEDA, *supra* note 73, § 1504, at 383, or "improve[s] competitive outcomes," Greaney, *supra* note 77, at 647.

Further, the intramarket second-best defense emphasizes dynamic efficiency concerns. Unlike Brodley, who would bifurcate the procedure into an ex ante and an ex post review, *see* Brodley, *supra* note 21, at 1048-49, my proposal would make all assessments at the initial proceeding. Also unlike Brodley, my concern is not in ensuring the eventual reemergence of market competition. *See id.* at 1037-38. Rather, my defense merely leaves the door open for subsequent public or private efforts to provide more effective remedies to the underlying market failure. Brodley's requirement of the reemergence of competition seems oddly misdirected when the underlying problem is one of market failure. Brodley's requirement stems from his view that antitrust law must safeguard individual consumer surplus and his belief that "competition" is the best method to ensure that consumers receive their fair share of the economic pie. *See id.* at 1038-39 ("This condition attempts to assure that consumers will eventually receive an appropriate share of the increased social wealth created by the subordination of their immediate interest. An appropriate share for consumers is simply the share of economic surplus that a competitive market would provide."). Many market failures are chronic rather than transitory. If a noncompetitive solution is both welfare-enhancing and stable over time, it should not be condemned merely because of its longevity. The policy objective should be to encourage the emergence of the private or public response that is most capable of improving social welfare, not the imposition of "competition" in settings where competition is unworkable or undesirable. This difference in market failure defenses once again illustrates how total welfare and consumer surplus orientations can lead to very different doctrinal standards.

Finally, the intramarket second-best defense advocated here differs from Leslie's defense in that it is framed in terms of a standard that can be used to assess private responses to market failures, rather than a set of economic characteristics that are likely to be present if such a standard is to be satisfied.

marketed that otherwise might not be produced.[93] Establishing the dynamic efficiency of such ventures would entail proving many of the same things that must be proven to show that the restraints are reasonably necessary to the venture and appropriately narrow in scope under the traditional ancillary restraints test: that the venture parties consist of a reasonably small percentage of all market actors, that the venture will cross-license new technology, that the venture is designed to ensure reasonable amounts of competition in the new product market, and that there are appropriate internal restrictions to prevent forms of spillover collusion.[94] Similar considerations would be rele-

93. I do not want to trivialize the issues involved in assessing the welfare effects of cooperative R&D. R&D joint ventures can raise difficult and complicated questions. Some of the theoretical difficulties were suggested earlier. *See supra* note 51 (examining welfare effects of cooperative R&D). Patents and intellectual property rights have also been the subject of extensive investigations, yielding conflicting results on questions such as what is the optimum length of time for patent protection, *see generally* W. NORDHAUS, INVENTION, GROWTH AND WELFARE 70-90 (1969); F.M. SCHERER, INDUSTRIAL MARKET STRUCTURE AND ECONOMIC PERFORMANCE 442 (2d ed. 1980); Richard Gilbert & Carl Shapiro, *Optimum Patent Length and Breadth*, 21 RAND J. ECON. 106 (1990); Edwin Mansfield, *Patents and Innovation: An Empirical Study*, 32 MANAGEMENT SCIENCE 175 (1986), and what are the most likely sources of invention in the economy, *see* GILBERT KIVENSON, THE ART AND SCIENCE OF INVENTING 2 (2d ed. 1982); TREVOR I. WILLIAMS, A SHORT HISTORY OF TWENTIETH CENTURY TECHNOLOGY 13 (1982). While these issues are difficult to answer in the abstract, they may be substantially easier to answer on a case-by-case basis, especially in light of the self-selection process involved in privately choosing research projects and initiating cooperative ventures. For further discussion of these issues, see generally Louis Kaplow, *The Patent-Antitrust Intersection: A Reappraisal*, 97 HARV. L. REV. 1813 (1984), and Robert P. Merges & Richard Nelson, *On the Complex Economics of Patent Scope*, 90 COLUM. L. REV. 839 (1990).

The social value of the cooperative R&D will vary significantly depending upon the type of research at issue. Basic research usually presents stronger public good justifications than research targeted at marginal advances in existing products, such as in-the-front-door ice water for the family refrigerator. The total welfare standard should be flexible enough to accommodate such differences.

94. R&D joint ventures are evaluated under the rule of reason. *See* National Cooperative Research Act of 1984, Pub. L. No. 98-462, 98 Stat. 1815 (codified at 15 U.S.C. §§ 4301-05 (1988)). The Act mandates rule-of-reason analysis and de-trebles damages for those ventures that comply with its registration provisions. The rule of reason as applied to R&D joint ventures invokes an abstract balancing of pro- and anticompetitive effects. This balancing often lacks specific analytic content. "The Department of Justice will not challenge a true joint research and development venture on antitrust grounds, unless the Department determines that the venture will have anticompetitive effects in a properly defined market that outweigh the venture's procompetitive benefits." *American Heart Association Pharmaceutical Roundtable*, 1998 DOJBRL LEXIS 7, *7-*8 (Mar. 20, 1998); *see also U.S. Dep't of Justice Antitrust Guidelines for International Operations*, *supra* note 49, at 20,627 ("[T]he possibility of anticompetitive effects would not lead the Department automatically to condemn the joint venture. The Department would consider whether the threat of anticompetitive effects would be outweighed by procompetitive efficiencies that the parties claimed would result from the joint venture.").

Greater insight can be gained by looking at the factors the Department of Justice deems relevant to its analysis. Although technically superseded by the 1945 *Antitrust Guidelines for the Licensing of Intellectual Property*, *supra* note 50, Case 6 of the earlier *Antitrust Guidelines for International Operations* still provides a useful illustration. No anticompetitive effects are thought to be likely if there are at least four other comparable R&D efforts under way in the same R&D market. *See id.* at 20,625. Importantly, the DOJ recognizes that co-

vant in establishing that the venture is the least restrictive means of obtaining the alleged benefits.

While the facts of *California Dental* are somewhat problematic, one can conceive of private restraints designed to counteract informational market failures that could withstand scrutiny under the test. The advertising restriction in *California Dental* is troublesome because closer examination may well demonstrate that the restraint is anticompetitive and will decrease total welfare. A more defensible example would be a private agreement by producers to standardize the reporting format of information for consumers. These types of agreements can facilitate consumer price comparisons and reduce consumer search costs.[95] Standardization agreements can also pose competitive risks. Standardized agreements can make it easier for com-

---

ordination in the market for the new technology may be essential to provide the necessary incentives to undertake the research in the first place. *See id.* at 20,626 ("Without such joint coordination, the value of the parties' R&D might be dissipated through competition in the product market. As a result, firms might avoid efficient joint R&D altogether, which might result in the costly duplication of R&D efforts. Coordination in markets using the technology output of the joint venture is therefore often essential to beneficial R&D."). Many procompetitive benefits can flow from cooperative R&D: the creation of new technologies, products, and processes that might not otherwise exist; the sharing of economic risk; pooling of complementary information, technology, and skills; low-cost transfer of information between participants; and the prevention of free-riding. *See id.* at 20,627. Anticompetitive threats are also real. The DOJ is concerned that venture restrictions may be a sham to coordinate collusion in existing product markets, or that the sharing of information about the joint venture can facilitate cooperation outside the venture — forms of spillover collusion. *See id.* at 20,626. The DOJ will review potentially anticompetitive aspects of the venture under the traditional test for ancillary restraints, examining whether they are "reasonably necessary to the successful operation of the venture." *Id.* at 20,627. R&D joint ventures raise difficult questions about membership. Overinclusive membership can undermine incentives to innovate. Underinclusive membership can potentially deny non-members technology that may become essential to remain competitive.

There is nothing in this analysis that necessarily conflicts with the affirmative defense proposed in this Article. Indeed, the framework outlined for evaluating intramarket second-best tradeoffs may add rigor and clarity to the evaluation of R&D joint ventures. The focus on market failures ensures that cooperative efforts are grounded in economic settings where cooperation is most likely to be necessary and desirable. Assessing the impact of the venture on total welfare provides a clearer metric for analysis than does the abstract balancing of pro- and anticompetitive effects, even if a number of the analytical factors are the same. Finally, the focus on dynamic efficiency makes explicit a set of competitive concerns that are implicit in the present analysis, but not systematically developed. Not surprisingly, defendants would most likely favor existing rule-of-reason evaluation. Clarity of analysis may or may not be a strategic virtue for litigants. More importantly, defendants would resist the heightened evidentiary burdens imposed by the rule advocated in this Article.

95. *See* Schwartz & Wilde, *supra* note 53, at 672-73 ("[T]he government should consider relaxing the antitrust laws to permit more voluntary standardization of the format in which contract prices and terms are quoted.... This is not to say that the government should benignly view agreements by firms to fix uniformly the substantive aspects of their transactions, but only to suggest the social desirability of agreements to use a common *format* to set out in a standard fashion the terms and prices each firm individually chooses to offer."); *see also* Richard Craswell, *Tying Requirements in Competitive Markets: The Consumer Protection Issues*, 62 B.U. L. REV. 661, 691 (1982) (discussing the benefits of standardized disclosure requirements to facilitate consumer price comparisons).

petitors to enter into collusive agreements, or may make the detection of cheating on cartel agreements easier to identify and privately enforce. Moreover, the agreement to a standardized format may spill over into agreements on substantive aspects of the terms or conditions of sale.[96]

In terms of the intramarket second-best defense, proponents of a private standardization agreement would first have to document the various market failures being addressed. Here, the focus would not simply be on asserting the existence of informational market failures, but on a detailed description of the nature of the goods in question, consumer search costs, and the buying practices of different purchasing groups.[97] Second, proponents would have to demonstrate a likely

96. Some of these concerns can be illustrated by the agreement in *Catalino, Inc. v. Target Sales, Inc.*, 605 F.2d 1097 (9th Cir. 1979), *rev'd*, 446 U.S. 643 (1980). The Ninth Circuit held that a horizontal agreement among beer wholesalers to eliminate short-term credit terms was not per se illegal, but would be evaluated under the rule of reason. The court reasoned that agreement might actually enhance competition. "[C]ompetition could be fostered by the increased visibility of price made possible by the agreement to eliminate credit. For example, an agreement to eliminate credit might foster competition by increasing the visibility of the price term, and hence, promote open price competition in an industry in which imperfect information shielded various sellers from vigorous competition." 605 F.2d at 1099. The court favorably cited the work of Lawrence Sullivan discussing the potential benefits of standardization. *See* 605 F.2d at 1099 n.5 (citing L. SULLIVAN, HANDBOOK OF THE LAW OF ANTITRUST § 99, at 277 (1977)). The Supreme Court reversed, holding that a horizontal agreement not to compete on credit terms was a per se illegal price fixing agreement. *See* 446 U.S. at 647. Significantly, the agreement in *Catalino* goes far beyond the types of format standardization advocated by Shwartz & Wilde, *supra* note 53. The beer wholesalers agreed to fix a substantive term of the agreement, presenting a clear competitive threat without any sufficiently persuasive offsetting welfare gain.

97. Unfortunately, the Supreme Court's tying cases create confusion as to when informational "market imperfections" can or should be recognized for antitrust purposes. In *Jefferson Parish Hospital District No. 2 v. Hyde*, the Court reasoned that while market imperfections "may generate 'market power' in some abstract sense, they do not justify the kind of market power that justifies condemning tying." 466 U.S. 2, 27 (1984); *see also* Town Sound & Custom Tops v. Chrysler Motors, 959 F.2d 468, 492-93 (3d Cir. 1992) (acknowledging the relevance of imperfect information and other market failures, but concluding that institutional considerations caution against using antitrust laws to correct a variety of "consumer protection" issues). In *Eastman Kodak Co. v. Image Technical Services, Inc.*, however, the Court was greatly influenced by the effects of information failures in the market for high-end photocopying equipment, and in the aftermarket for Kodak parts and service. "For the service-market price to affect equipment demand, consumers must inform themselves of the total cost of the 'package' — equipment, service, and parts — at the time of purchase; that is consumers must engage in accurate lifecyle pricing." 504 U.S. 451, 473 (1992). The inability to engage in life-cycle pricing was critical to the Court's finding of market power on the part of Kodak. It is difficult to reconcile the Court's teachings in *Jefferson Parish* and *Kodak*. *See* Michael S. Jacobs, *Market Power Through Imperfect Information: The Staggering Implications of* Eastman Kodak Co. v. Image Technical Services *and a Modest Proposal for Limiting Them*, 52 MD. L. REV. 336, 357 (1993) ("The approaches of *Kodak* and *Jefferson Parish* to the role of informational imperfections in market power analysis cannot be reconciled."). *But see* Little Caesar Enters., Inc. v. Smith, 34 F. Supp. 2d 459, 484-85 (attempting to reconcile *Jefferson Parish* and *Kodak* in terms of determining when imperfect information can generate market power for tying purposes by focusing on whether the informational difficulties were exacerbated by the entity charged with having market power).

increase in total welfare (static efficiency). Standard reporting formats can make markets function more effectively by decreasing consumer search costs and increasing competition between producers. The result may be reductions in price or increased levels of quality and service. These welfare gains would have to be balanced against anticompetitive risks, such as the increased risk of collusion. Here, the concern is making sure that the agreement is focused on the format for conveying information and not on substantive terms of the transaction. There do not appear to be serious dynamic efficiency concerns, although judges should be aware of the possibility that disclosure requirements can potentially create or increase barriers of entry. Additionally, disclosure requirements are not always benign. There is often a tendency for competitive energies to be channeled into the categories being reported to consumers. Finally, courts should inquire into the possibility of achieving similar welfare benefits through less restrictive means.

As one moves from cases that can fit under existing rule-of-reason analysis to cases that would require express doctrinal recognition of an intramarket second-best defense — typically cases involving a more direct challenge to established presumptions of anticompetitive effects — one can envision similar types of economic theories and empirical evidence that could be used to satisfy the requirements of the affirmative defense. The more novel the claim, however, the more skeptical the factfinder should be in considering the arguments.

The making of an intramarket second-best defense is perhaps most clearly illustrated in the context of negative externalities. Assume that defendants are cigarette manufacturers charged with forming a cartel to fix the price of cigarettes (or potentially agreeing to a merger that violates section 7 of the Clayton Act).[98] A prima facie antitrust viola-

---

The affirmative defense for intramarket second-best tradeoffs avoids a number of these problems. Under a total welfare standard, the market failures acknowledged in *Jefferson Parish* and *Kodak* would both be relevant, but the inquiry would focus on the relationship between the informational market failure, the alleged restraint, and total welfare. The difficulties the courts are having conducting rule-of-reason analysis in markets characterized by imperfect information suggest some of the limitations of existing doctrine, which focuses on an abstract balancing of pro- and anticompetitive effects. They also suggest that rule-of-reason analysis might be improved if it adopted portions of the framework developed here to examine intramarket second-best tradeoffs. Is the restraint responsive to an identifiable market failure? Will the conduct increase total welfare? Are there dynamic efficiency concerns? Is there a less restrictive alternative?

98. Some people may object to the cigarette example because it permits tobacco companies that are already causing substantial public health harms to benefit from the windfall of monopoly profits. The argument in favor of taxing away such cigarette profits may indeed be strong. Nevertheless, the deeper insight of being able to use market structure as a means of correcting negative externalities should not be lost. One can imagine cases where permitting producers to earn monopoly rents as a means of correcting externalities would not strike people as morally objectionable. For example, assume two timber companies are logging a national forest. For political and other reasons, they are not paying the full social costs of their logging, which means that their customers are buying lumber "too cheaply." If they

tion would be established by documenting the existence of an agreement to fix prices (or defining the market and documenting concentration ratios that trigger a presumption of illegality). The defendants would be able to rebut the presumption of a merger's illegality if they were able to demonstrate with clear and convincing evidence the following contentions. First, the existence of a market failure — cigarettes impose substantial negative externalities because the social costs of smoking (increased Medicare and Medicaid expenses, harm of second-hand smoke, etc.) exceed the private cost of cigarettes to consumers as reflected in the prevailing market price, including government taxes. Second, the cartel price (or the increased market price associated with increased market power from the merger) would result in a net increase in total welfare (static efficiency). This argument could be made with expert economic testimony employing arguments and analysis similar to those explored in Section I.B.1 and illustrated in figure 2 — the higher price associated with the otherwise anticompetitive conduct moves society closer to the "price equals marginal social cost" equilibrium.[99]

Third, the defendants would have to prove that the cartel (or merger) would not structurally impair the ability of private or public actors subsequently to ameliorate the effects of the market failure (dynamic efficiency).[100] The concern over dynamic efficiency is a safe-

---

merged and attained some degree of market power, the price of their lumber goes up, production falls, negative externalities are diminished, and customers are paying closer to the full social costs of what they buy. Permitting the lumber companies to retain the profits does not strike me as morally objectionable, particularly in light of the increase in total welfare. This example also illustrates why a total welfare standard is not only more efficient, but may also be more fair than a consumer surplus standard, at least in cases of negative externalities. After the merger, the price of lumber is raised from a level that is "too cheap" to a level that more closely approximates the social costs of the timber. This reduces consumer surplus, but on these facts the consumer surplus represents the ill-gotten gains of over-exploiting the forests.

99. *See* Gulbrandsen & Skeath, *supra* note 31, at 9-15 (calculating the welfare effects of different market structures given the presence of negative externalities). The authors also consider the welfare effects of the government's 1911 antitrust breakup of the American Tobacco Company. Under monopoly conditions, the price of cigarettes was higher and output was lower than under competition. After the 1911 breakup, output increased. "In the two years following the 1911 breakup, output nearly doubled from its 1909 level." *Id.* at 4. The authors conclude that "it is certainly possible that leaving the American Tobacco Company untouched by antitrust action *may* have been in the best interest of American consumers and their families." *Id.* at 20 (emphasis in original).

100. Dynamic efficiency concerns will be difficult to operationalize. The third element calls upon courts to engage in forms of comparative institutional analysis. Both Greaney and Brodley acknowledge similar concerns in their advocacy of a market failure defense. *See* Brodley, *supra* note 21, at 1047-48 ("The existence of market failure indicates the need for regulation, but does not resolve whether the regulation should be public or private."); Greaney, *supra* note 77, at 630 ("This precedent militates in favor of conditioning the market failure defense on a showing that private ordering is superior to public regulation."). My intramarket second-best defense favors concurrent and competing public and private efforts to address market failures. The evaluation of dynamic efficiency is primarily concerned with

guard against anticompetitive conduct that produces a short-term increase in total welfare but that prevents more effective initiatives in the future.[101] In the cigarette example, such structural concerns are unlikely. A cartel is not likely to impede the ability of federal or state governments to impose future taxes that would compel smokers to internalize the effects of the externalities.[102] Indeed, on these facts, the disclosure required by the first element of the defense is likely to invite government action.[103] Finally, the defendants would have to prove that there is not a less restrictive private alternative to the restraint. The focus here is on the actions more consistent with the antitrust laws that the defendants could undertake to address the market failure. For example, under certain circumstances, a joint venture might be less restrictive than a complete merger. On these facts, a price fixing agreement may be less restrictive than a merger.

Hospital nonprice competition provides another useful illustration. Defendant hospitals would first have to document the many market failures that frustrate active price competition and encourage various forms of nonprice competition.[104] Second, the hospitals would have to prove that merging would result in a net increase in social welfare

---

avoiding private action that will substantially increase the cost of subsequent public or private responses to the underlying market failures.

101. Many unusual situations can be envisioned in assessing dynamic efficiency. Courts may be faced with competing private initiatives to correct the same or related market failures. In the tradition of second-best analysis, the welfare effects of these initiatives may vary depending upon whether they are viewed in isolation or in combination. Similarly, courts may be confronted with serial private efforts of an identical nature that may have different welfare effects. Some of these issues will be difficult to resolve in practice. My intention here, however, is to present a skeletal outline of the affirmative defense. The particular details of the analysis can be fleshed out as concrete problems arise. The rubric of "dynamic efficiency," however, should be flexible enough to permit courts to evaluate both actual and hypothetical competing public and private initiatives in whatever permutations seem appropriate.

102. It is conceivable that a merger or the formation of a cartel could facilitate private actors' ability to block future innovations such as the introduction of safer low-tar or low-nicotine alternatives. At the same time, the higher prices attributable to increased market power are likely to make smoking alternatives (cessation programs or safer cigarettes) relatively more attractive to some consumers. Also, in industries like cigarettes, one may need to be concerned about the effects that branding and advertising may have on subsequent entry into the market.

103. While it is possible that the judge's decision in a particular case could itself influence the likelihood of responsive political action, either in a positive or negative direction, I do not view this as an appropriate consideration under dynamic efficiency analysis. The focus should remain on a structural evaluation of whether the private conduct being approved would make it more costly to implement a public alternative in the future. While not part of the judicial analysis, the potential political backlash associated with the types of public disclosures required by the first element — admitting the negative and often deadly effects of smoking — is likely to serve as a serious deterrent to cigarette companies invoking the defense.

104. *See* Greaney, *supra* note 77 (discussing health care market failures); *see also supra* text accompanying notes 45-46 (discussing effects of market failures on price and nonprice competition).

(static efficiency). Unfortunately, unlike the case of the negative price externalities, there is no simple diagram that illustrates the costs of nonprice competition. Defendants could rely upon cross-sectional studies of the effects of market power and economic concentration upon medical expenses.[105] Additionally, the hospitals could introduce evidence that the costs of providing nonprice amenities exceed the social value of the attribute as measured by the individual consumer's willingness to pay for such unbundled options.[106] If the hospitals could show that reductions in competition increase total welfare and result in a more efficient distribution of resources, then the second element would be satisfied.

Third, the hospitals would have to show that a merger would not structurally undermine the ability of public or private actors subsequently to address the underlying market failures (dynamic efficiency). There are legitimate dynamic efficiency concerns in hospital markets. Mergers both increase the level of economic concentration and decrease the number of actors in the market. As such, hospital mergers threaten to undermine two alternative responses to the market failures triggering nonprice competition. The first response is the private restructuring of the firms providing health care — integrating the functions of financing and delivery — as witnessed by the rise of managed care plans.[107] The second response is the introduction of greater levels of bidding and selective contracting by large employers and by government-sponsored Medicare and Medicaid programs.[108] Both of these responses could be undermined by substantial increases in economic concentration at the hospital level. Thus, permitting the merger of hospitals in traditionally structured markets in order to counteract the effects of nonprice competition provides an example of conduct that may be statically efficient (resulting in a short-term increase in total welfare) but is dynamically inefficient (structurally undermining the ability of public and private actors to address subsequently the underlying market failures).[109] So to establish the defense,

105. *See* studies cited *supra* note 44 (empirical studies of the effects of competition in health care markets).

106. *See* discussion *supra* note 43 and accompanying text (using consumer willingness to pay for a series of unbundled options as a benchmark for assessing the social value of nonprice attributes).

107. *See* Hammer, *supra* note 45 (examining the changing structure of health care markets in terms of a Coasian transformation of the firms providing health care services and financing).

108. *See, e.g.*, Thomas L. Greaney, *Competitive Reforms in Health Care: The Vulnerable Revolution*, 5 YALE J. ON REG. 179, 185-87 (1988) (discussing the role of selective contracting in overcoming health care market failures).

109. *See* Hammer, *supra* note 45 (arguing that dynamic efficiency concerns should lead to the rejection of hospital mergers in those markets capable of making a successful transition from the Old to the Emerging Regime).

merging hospitals would have to show that these solutions were unlikely to emerge in their market in the foreseeable future.

Finally, to complete their intramarket second-best defense, the hospitals would have to demonstrate that there was not a less restrictive alternative to merger. Joint ventures for specific new technologies and the conscious splitting of service menus or allocation of the market may be less restrictive strategies.

An affirmative defense for intramarket second-best tradeoffs raises unsettled doctrinal questions and practical concerns. Practical concerns, such as the ability of courts to litigate second-best tradeoffs, are examined in Part II. The doctrinal issues associated with second-best tradeoffs and a total welfare standard are examined in Part III.

II. THE VIABILITY OF AN INTRAMARKET SECOND-BEST DEFENSE: THEORETICAL INTEGRITY, ABILITY TO LITIGATE, AND COST-BENEFIT ANALYSIS

It is one thing to assert that courts should assess intramarket second-best tradeoffs in light of a total welfare standard as a theoretical matter. It is quite another matter to contend that courts are capable of implementing such an approach in the context of litigation. The affirmative defense proposed in this Article incorporates a total welfare standard far more expressly than is currently recognized in antitrust law. To establish the viability of a total welfare standard, a number of propositions must be defended. Economic theories of consumer surplus, producer surplus, and total welfare must be able to provide a sound basis for decisionmaking. Moreover, it must be possible to translate these economic theories into the legal setting, by providing both the technical information and the intuitive framework necessary for deciding cases.

A. *The Underpinnings of Total Welfare: Understanding Consumer Surplus*

Consumer surplus plays a central role in welfare analysis.[110] At its most basic level, individual consumer surplus is the monetary difference between the amount a person pays for a good or service and the maximum amount he would be willing to pay for that same good or service.[111] For example, if I send an agent to bid at auction for a rare

110. Economists reading this Article may wish to skip to *infra* notes 121-122 and accompanying text, for what follows is a discussion of some basic economic principles for the general reader.

111. *See* FRANK CAMM, CONSUMER SURPLUS, DEMAND FUNCTIONS, AND POLICY ANALYSIS 1 (1983) ("Consumer Surplus is a monetary measure of the difference between what an individual pays for consuming a good or service and the amount he is willing to pay,

Cambodian sculpture and authorize my agent to bid up to a level that reflects my valuation of the art work, my consumer surplus would be the difference between the amount of my winning bid (assuming I am successful in my efforts) and the higher amount I had authorized my agent to bid (my willingness to pay). Consumer surplus at a market level is simply the aggregation of individual consumer surpluses. We can envision this (subject to a number of caveats) as the area between the market demand curve and the prevailing market price for any given commodity (see figure 1).[112] This is essentially the approach adopted by Alfred Marshall who popularized the concept in his 1890 text *Principles of Economics*.[113]

Policymakers are frequently concerned about the impact of different actions on consumer surplus. Take, for example, an increase in price. Marshall, in calculating the change in consumer surplus, holds the individual's income constant, assuming that the change in relative prices has no income effects. At the higher price, the individual consumes less of the good. This lower level of consumption reduces the individual's utility because he presumably gains benefits from having more of the good, for otherwise he would not have purchased it, whatever its price. The change in consumer surplus is the difference between benefits (utility) the consumer enjoyed before and after the price increase. Measuring those benefits in monetary terms, the difference can be approximated by the area to the left of the Marshallian demand curve between the two price levels.[114]

---

given his income and the prices he faces."). *See generally* E.J. MISHAN, WHAT POLITICAL ECONOMY IS ALL ABOUT: AN EXPOSITION AND CRITIQUE 69-82 (1982).

112. Camm notes that the characterization of consumer surplus as the "area under the demand curve" is appropriate "only if (a) the good in question has no close substitutes or complements or (b) consumer prices for all other goods are fixed." CAMM, *supra* note 111, at vii. "For practical purposes, *consumer surplus is not the area under the demand curve: It is the sum of areas to the left of the consumption loci.*" *Id.* Another difficulty in characterizing consumer surplus as the area under the demand curve is that different measures of consumer surplus envision the usage of different types of demand functions. *See id.* at 28-29.

113. ALFRED MARSHALL, PRINCIPLES OF ECONOMICS 175, 429, 446-55, 465-72 (1890).

114. Changes in individual utility can be translated into exact monetary values via the demand function, given Marshall's method of deriving consumer demand directly from consumer utility. *See* Peter C. Dooley, *Consumer's Surplus: Marshall and His Critics*, 6 CANADIAN J. ECON. 26, 27 (1983). Marshallian demand curves and his analysis of consumer surplus can be controversial. Dooley summarizes the major categories of criticisms that have been levied against Marshall:

> [F]irst, whether an additive utility function adequately explains consumer behavior; second, whether the marginal utility of money can be treated as a constant; third, whether the quantity demanded of one commodity can be treated as a function of its price alone; and fourth, whether it is possible to make interpersonal comparisons.

Id. at 28. He elaborates on each of these critiques. *See id.* at 28-35; *see also* Robert B. Ekelund & Robert F. Hebert, *Consumer Surplus: The First Hundred Years*, 17 HIST. POL. ECON. 419, 435-39, 445-47 (1985) (discussing historical and modern critiques of Marshall's theory of consumer surplus); Edward R. Morey, *Confuser Surplus*, 74 AM. ECON. REV. 163

A serious limitation of Marshall's analysis is his assumption that price changes have no income effects. Empirically, this assumption is most problematic when the good in question accounts for a large share of an individual's expenditures, or when consumption of the good itself is highly sensitive to fluctuations in personal income. Hicksian measures of consumer surplus attempt to address these considerations. Hicks introduced two alternative measures to assess changes in consumer surplus: the Hicksian compensating variation and the Hicksian equivalent variation.[115] The compensating variation asks, when there is a price increase, how much money would be required to return the individual to his pre-price-increase level of utility.[116] The equivalent variation asks what reduction in income would be required to reduce the individual's utility to the same degree as the increase in price.[117] Phrased slightly differently, the compensating variation asks what monetary transfer would be required to compensate fully the individual for the effects of the price increase and return him to his original level of utility (albeit in a world with higher prices). The equivalent variation asks how much the individual would be willing to pay in order to avoid the effects of the price increase (prices are unchanged in this variation, but the individual loses income and hence utility).[118] While closely related in approach, the compensating and the equivalent variations will usually produce different monetary measures of changes in consumer surplus.[119] Moreover, the Hicksian and Marshallian measures will yield different assessments of the impact of various policies on consumer surplus.[120]

Although subject to numerous technical refinements and professional debate, the notion of consumer surplus possesses a relatively simple and intuitive core. At its heart, Marshall's consumer surplus is the difference between the actual price charged and an individual's

(1984) (questioning whether dollar measures of individual utility exist that would permit a measure of preference intensity and interpersonal comparisons).

115. For expositions of the Hicksian measures of consumer surplus, see CAMM, *supra* note 111, at 6-7, and Ekelund & Hebert, *supra* note 114, at 440-43. *See also* KREPS, *supra* note 5, at 58-59; VARIAN, *supra* note 4, § 15.4.

116. *See* CAMM, *supra* note 111, at 7 ("[W]e can ask how much the individual's income would have to be raised following the price change to achieve the level of utility he enjoyed before the price change.").

117. *See id.* ("[W]e can ask what cut in income will hurt the individual as much as the price change does.").

118. *See id.* at 9-10 (graphic and textual illustrations).

119. *See id.* at 25-26 n.2 ("When the price of a normal good changes, compensating variation exceed equivalent variation if the price raises; the reverse holds if price falls.").

120. For comparisons of the Hicksian and Marshallian approaches, see *id.* at 28-29, and Jerry A. Hausman, *Exact Consumer's Surplus and Deadweight Loss*, 71 AM. ECON. REV. 662 (1981). For an illustration of the differences these approaches yield in measuring the deadweight loss, see CAMM, *supra* note 111, at 63-64, and Hausman, *supra*, at 672-73.

willingness to pay. More refined Hicksian notions of consumer surplus also invite introspection in which most people can engage. How much money would be necessary to undo the harm of an adverse event, such as an increase in prices, or how much money would an individual be willing to pay to avoid a price increase (or other adverse event)?

These concepts resonate closely with common law calculations that ordinary jurors perform on a routine basis. Expectation damages in contract cases are designed to protect an individual's benefit of the bargain.[121] In calculating and awarding expectation damages, jurors are essentially awarding the nonbreaching party their "surplus" associated with the contract. Similarly, compensatory damages in torts perform the same basic calculation envisioned by Hicksian compensating variations. The objective is to provide damages (a money award) that attempt to return the plaintiff to the same level of well-being (utility) that he enjoyed before the tort. These calculations are not necessarily easy to make, nor does the judicial system always get them right, but the common law analogies tell us that the calculations can be reduced to intuitions ordinary jurors and judges can use with levels of accuracy and effectiveness deemed acceptable in other important areas of law.[122] Additional examples of relatively complicated economic assessments being operationalized at an intuitive level can be found in the assessment of negligence in tort cases, the reasonable person standard, and forms of cost-benefit analysis or balancing tests used throughout the law. I point these out because they tend to ne-

121. The example of expectation damages focuses on individual rather than aggregate consumer surplus. My point, however, is not that ordinary citizens can be expected to calculate aggregate consumer surplus, but rather that the central notions can be made intuitive and individually accessible. Intuitive juror understanding will not be a substitute for expert economic testimony, but it means that properly done, there is an effective foundation for the economist's testimony and the lawyer's argument.

122. To be persuasive, the common law analogy to torts and contracts must also examine whether there are any acceptable institutional substitutes for, or superior alternatives to, these common law regimes. The ability to litigate a standard is a necessary but insufficient condition for adopting a standard. "Can" and "should" raise separate issues. Whether a standard "should" be litigated raises questions of comparative institutional analysis. With market failures and intramarket second-best tradeoffs, for example, regulatory substitutes may be available that are not available for torts and contracts. Furthermore, administrative agencies may have a comparative advantage in determining consumer surplus and total welfare. Consequently, it might make sense to entrust juries with the calculation of tort and contract damages, but not to entrust antitrust juries with calculating consumer surplus. While the present Section focuses on whether consumer surplus "can" be operationalized, the question of whether it "should" be adopted is strengthened by the realization that antitrust law supplies default rules for private markets that can be trumped by federal and state regulatory initiatives if regulation has a strong comparative advantage. Moreover, the third element of the affirmative defense (dynamic efficiency), outlined *supra* Section I.D, seeks to permit private action only in those cases in which it would not substantially impair the ability of subsequent efforts to address the underlying market failures. I am grateful to Louis Kaplow for his comments highlighting potential differences between common law and regulatory assessments in this area.

gate an obvious objection to my scheme: namely, that jurors and judges will be unable to make the calculations it calls for. While these calculations are not the same as the ones in these examples, they relate to notions of personal and economic well-being and are often at a similar level of complexity.

B. *Theory into Practice: Litigating Second-Best Claims*

While intuitively grounded, theories of consumer surplus are also capable of being operationalized in a manner that can facilitate technical decisionmaking. Econometricians used to be faced with a dilemma. Marshallian demand is observable and may be roughly calculated from information on actual consumer purchases at different price levels.[123] Hicksian compensated demand curves, which attempt to account for the income effects of price changes, are not observable,[124] yet it is generally believed that Hicksian demand curves provide more defensible estimates of changes in consumer surplus.[125] This is because they incorporate the income effects that are associated with the change in relative prices. The old dilemma was between use of an easier to estimate but less accurate Marshallian measure, or a more difficult to calculate but more accurate Hicksian measure.

Fortunately, this dilemma no longer is a serious obstacle in most cases. Robert Willig has demonstrated that calculations of Marshallian consumer surplus based upon market demand functions often provide good approximations of either Hicksian measure.[126] More im-

123. *See* CAMM, *supra* note 111, at 37 ("Only one of these — the Marshallian — is a behavioral function in the sense that it describes actual behavior in the absence of the forms of compensation envisioned in the other measures.").

124. *See id.* at 64 ("Hick's . . . functions are hypothetical functions that must be inferred from the Marshallian functions we can observe directly in the absence of compensation.").

125. In his treatment of the efficiency defense, Williamson relied upon Marshallian demand curves and Marshall's notion of consumer surplus. *See* Williamson, *Economies as an Antitrust Defense*, *supra* note 66, at 22 n.4; Williamson, *Economies Revisited*, *supra* note 66, at 708 n.28. Fisher & Lande subsequently levied the standard critique that, for "accurate measurement of deadweight losses and consumer/producer surplus, one must perform welfare analysis using income-compensated, or Hicksian, demand curves." Fisher & Lande, *supra* note 23, at 1629 n.176.

126. *See* Robert Willig, *Consumer's Surplus Without Apology*, 66 AM. ECON REV. 589 (1976); *see also* CAMM, *supra* note 111, at 62 ("Willig has quantified the divergence of measures based on the Marshallian and Hicksian curves in terms of income elasticities and factor shares for arbitrary functional forms. Where these numbers are not large, his technique may prove useful in moving from one measure of consumer surplus to another."); Hausman, *supra* note 120, at 662 ("Robert Willig derives bounds for the percentage difference between the correct measure of either the compensating or equivalent variation and the Marshallian measure derived from the market demand curve. His bounds . . . demonstrate that the Marshallian consumer's surplus is often a good approximation to Hicks' consumer's surplus. . . . Thus he hopes to remove the need for apology that applied economists often need to give to theorists who remark on the inappropriateness of using Marshallian consumer's surplus to measure welfare change.").

portantly, Jerry Hausman has illustrated a technique by which it is possible to derive Hicksian compensated demand functions from observable market demand and to directly estimate Hicksian compensated and equivalent variations.[127] Thus, theoretical and applied economics now provide a set of tools that make it possible to litigate issues such as the likely effect of changes in market structure, or conduct on consumer prices and hence consumer surplus.[128] Obviously, consumer surplus constitutes only one part of the total welfare puzzle, but it is an essential component that must fall into place if antitrust law is to move in the direction of adopting a total welfare orientation.

Recent developments in the assessment of the unilateral effects of merger in markets with differentiated products further illustrate the ability of economists to operationalize sophisticated antitrust theories that require estimations of actual consumer demand.[129] At least since the 1992 Joint Federal Trade Commission and Department of Justice Merger Guidelines, antitrust enforcement agencies have treated seriously the possibility of anticompetitive effects flowing unilaterally from mergers.[130] Unilateral effects are anticompetitive effects that can

---

127. *See* Hausman, *supra* note 120, at 663 ("The basic idea used in deriving the exact measure of consumer's surplus is to use the *observed* market demand curve to derive the *unobserved* compensated demand curve. It is this latter demand curve which leads to the compensating variation and the equivalent variation."); *see also* Robert H. Haveman et al., *Exact Consumer's Surplus and Deadweight Loss: A Correction*, 77 AM. ECON. REV. 494 (1987).

128. *See* CAMM, *supra* note 111, at 1 ("The current theoretical literature on consumer surplus and social welfare measures seems to be reaching a consensus that well defined ('path-independent') measures of changes in consumer surplus based on Hicksian income-compensated demand functions can be developed strictly as functions of observable data on quantities and prices.").

129. *See* Jonathan B. Baker, *Unilateral Competitive Effects Theories in Merger Analysis*, ANTITRUST, Spring 1997, at 21; Thomas Overstreet et al., *Understanding Econometric Analysis of the Price Effects of Mergers Involving Differentiated Products*, ANTITRUST, Summer 1996, at 30; Carl Shapiro, *Mergers with Differentiated Products*, ANTITRUST, Spring 1996, at 23; Christopher A. Vellturo, *Evaluating Mergers with Differentiated Products*, ANTITRUST, Spring 1997, at 16; Gregory J. Werden, *Simulating Unilateral Competitive Effects from Differentiated Products Mergers*, ANTITRUST, Spring 1997, at 27. For more technical assessments of the economic issues raised by unilateral market effects, see Jonathan Baker & Timothy F. Bresnahan, *Estimating the Residual Demand Curve Facing a Single Firm*, 6 INT'L J. INDUS. ORG. 283 (1988); Jerry A. Hausman & Gregory K. Leonard, *Economic Analysis of Differentiated Products Mergers Using Real World Data*, 5 GEO. MASON L. REV. 321 (1997); Jerry Hausman et al., *Competitive Analysis with Differentiated Products*, 34 ANNALES D'ECONOMIE ET DE STATISTIQUE 159 (1994); Jerry A. Hausman et al., *A Proposed Method for Analyzing Competition Among Differentiated Products*, 60 ANTITRUST L.J. 889 (1992); Gregory J. Werden, *Simulating Effects of Differentiated Products Mergers: A Practical Alternative to Structural Merger Policy*, 5 GEO. MASON L. REV. 363 (1997); Gregory J. Werden, *A Robust Test for Consumer Welfare Enhancing Mergers Among Sellers of Differentiated Products*, 44 J. INDUS. ECON. 409 (1996); and Gregory J. Werden & Luke M. Froeb, *The Effects of Merger in Differentiated Products and Industries: Logit Demand and Merger Policy*, 10 J.L. ECON. & ORG. 407 (1994).

130. *See U.S. Dep't of Justice & U.S. Fed. Trade Comm'n Horizontal Merger Guidelines*, 4 Trade Reg. Rep. (CCH) ¶ 13,104 (§ 2.211) (Apr. 7, 1992) [hereinafter *DOJ and FTC Merger Guidelines*]. The examination of unilateral effects can be contrasted with the tradi-

occur without any combination or collusion with nonmerging parties. Adverse unilateral effects are most likely to occur in markets with differentiated products where the merging parties' products are perceived to be the first and second choices for a significant segment of consumers. If Product A (let us call it M&Ms) and Product B (let us call it Reese's Pieces) are a consumer's first and second choices, and each are independently produced (by Mars and Hershey), one of the constraints Mars faces in deciding whether to charge more for M&Ms is fear of lost sales (and profits) to Hershey, which will sell more Reese's Pieces.[131] The fraction of M&Ms sales that are lost to Reese's Pieces is defined as the "diversion ratio."[132] If Mars and Hershey merge, however, the diverted sales are not "lost" to Mars-Hershey and no longer serve to deter potential price increases for M&Ms.

One approach to estimating the adverse unilateral effects of a merger is to specify a full system of demand and supply equations and to measure the "own" and "cross" elasticities for all of the relevant products and producers.[133] The ability to specify the appropriate demand and supply equations depends on the availability and adequacy of the underlying data. In practice, such data are becoming increasingly available with improved computerized records and machine-scanned retail sales.[134] With these data in hand, numerous techniques are available for economists to specify the appropriate demand system.[135] Once a system of equations is specified, enforcement agency

---

tional merger analysis of coordinated effects reflected in the problem of collusion. Unilateral effects can occur in the absence of any collusion and under the assumption that all members of the industry continue to act in an independent and individually profit-maximizing fashion.

131. *See* Shapiro, *supra* note 129, at 24-25; Vellturo, *supra* note 129, at 16-17.

132. *See* Vellturo, *supra* note 129, at 17 ("The Diversion Ratio of Product A to Product B is defined as the portion of sales the producer of Product A would lose on Product A as a result of a price increase on Product A . . . .").

133. *See* Overstreet et al., *supra* note 129, at 31. Vellturo terms this type of approach "structural demand analysis." Vellturo, *supra* note 129, at 16.

134. *See* Shapiro, *supra* note 129, at 25; *see also* Hausman & Leonard, *supra* note 129, at 325 (describing how Nielsen retail scanner data can be used to estimate consumer demand and calculate own and cross price elasticities of various products). The systematic collection and subsequent sampling of retail scanner data makes it possible to obtain fairly accurate estimates of consumer demand for those products sold and tracked with bar codes. *See id.*

135. Vellturo summarizes some of the approaches that have been taken to "structural demand analysis":

> The approaches to structural demand analysis include the AIDS Model (Hausman, Leonard & Zona), the Antitrust Logit Simulation Model (Froeb & Werden), and the Residual Demand Elasticity Model (Baker). All three methods within this line provide sufficient structure to estimate (or, at least, simulate) the price effect of the proposed transaction . . . . All three are also highly data-intensive: the models require detailed price and sales data for the merging products (AIDS Model); price and sales data on the merging products and detailed industry-specific and product-specific cost data (Residual Demand Elasticity Model); or market share, and pre-determined elasticity estimates (Antitrust Logit Model).

economists can, and now do, run simulations to estimate the likely price effects of a merger.[136] The task is complicated. In addition to specifying the supply and demand equations and calculating the appropriate diversion ratios, to do a complete analysis it is necessary to formally model and evaluate likely competitor reactions in the form of new entry or the repositioning of product lines, as well as to factor in any likely efficiency gains that could change the cost structure of the merging firms and increase productive efficiency. But this is also now being done.[137] It is a short step from completing this type of analysis to assessing the impact of merger on consumer surplus and total welfare.[138]

Vellturo, *supra* note 129, at 16 (referring to Hausman et al., *Competitive Analysis with Differentiated Products*, *supra* note 129 (AIDS Model), Werden & Froeb, *supra* note 129 (Antitrust Logit Model), and Baker & Bresnahan, *supra* note 129 (Residual Demand Model)). For further discussion of each of these techniques, see Overstreet et al., *supra* note 129, at 31-33.

This literature includes a handful of examples where the authors apply their models to particular markets to assess the likely unilateral effects of mergers. *See* Baker & Bresnahan, *supra* note 129, at 290-98 (calculating the residual demand elasticities for three producers of domestic beer — Anheuser-Busch, Pabst, and Coors); Hausman et al., *Competitive Analysis with Differentiated Products*, *supra* note 129, at 162-72 (illustration estimating consumer demand for beer and calculating the own and cross price elasticities for various brands); Hausman & Leonard, *supra* note 129, at 335-36 (examining the market for bath tissue, estimating consumer demand, calculating own and cross price elasticities, and predicting the unilateral price effects of a merger between Kimberley-Clark and Scott); Werden & Froeb, *supra* note 129, at 417-20 (simulating the welfare effects of various hypothetical mergers amongst U.S. long distance carriers using the antitrust logit model, data of the relevant market shares, and estimated demand elasticities).

136. There are less technical and less data-intensive approaches to assessing the likely unilateral effects of merger. These methods center on obtaining accurate estimates of the diversion ratio as a means of determining the degree to which the pre-merged parties served as deterrents to unilateral price increases. These techniques include switching studies, win/loss reports, and end user surveys. *See* Shapiro, *supra* note 129, at 25; Vellturo, *supra* note 129, at 18-19. These techniques can be employed when data or technical restrictions prevent full structural demand analysis.

137. Shapiro outlines a four step process in unilateral effects analysis: (1) calculation of the diversion ratio; (2) estimation of the likely post-merger price increase assuming no reactions by competitors and no synergies to the merging parties; (3) accounting for the likely strategic reaction of competitors; and (4) accounting for any likely cost-reducing effects of the merger. *See* Shapiro, *supra* note 129, at 24. The primary forms of strategic reaction by competitors consist of new entry into the market or the repositioning of products by existing competitors. *See id.* at 27-28. An assessment of synergies would include any cost savings or efficiencies attributable to the merger. *See id.* at 28.

138. Werden outlines the steps of simulating the effects of mergers and assessing unilateral price effects. These steps include calibrating the demand system, estimating marginal costs, and calculating post-merger prices and outputs. *See* Werden, *Simulating Unilateral Competitive Effects from Differentiated Products Mergers*, *supra* note 129, at 28. Werden also corrects for entry and efficiency concerns. *See id.* All of these tools permit the enforcement agencies to assess the impact of the proposed merger on total welfare.

> One reason for undertaking merger simulation is that it permits a quantitative trade-off, or netting out, of efficiencies against the price-increasing effects of internalizing competition between the merging firms.... It is also possible to determine the net effect of a merger on total economic welfare, which includes the profits of the firms.

Id.

It is still too early to assess the impact that the economic theories and econometric techniques underlying the unilateral effects analysis will have on antitrust enforcement, counseling, and litigation. To date, most such analyses have been employed internally, by enforcement agencies deciding whether to permit or challenge transactions. Testimony reporting similar work has, however, been used in litigation and relied upon by courts in reviewing the legality of mergers.[139] While the doctrinal issues raised by unilateral effects are substantially less threatening to traditional antitrust approaches than those posed by intramarket second-best tradeoffs, many of the tools used in assessing the claims are mostly similar, both in terms of economic substance and in terms of the complexity of the analysis.

Two preliminary lessons can be drawn. First, the economic tools and legal sophistication required to litigate claims predicated on specifying complete systems of supply and demand equations and estimat-

---

139. *See* Shapiro, *supra* note 129, at 29 (citing New York v. Kraft Gen. Foods, Inc., 926 F. Supp. 321, 1995-1 Trade Cas. (CCH) ¶ 70,911 (S.D.N.Y. 1995), and United States v. Gillete, 828 F. Supp. 78 (D.D.C. 1993)). The district courts in both *Kraft* and *Gillette* rejected the unilateral effects arguments, not as a doctrinal matter, but on the facts before the court. In *Kraft*, the State of New York challenged Kraft's acquisition of Nabisco's ready-to-eat breakfast cereal assets. An important part of the State's case was the alleged unilateral price effects the merger would have on specific brands. In particular, the State maintained that the acquisition would permit Kraft to raise the price of its Post Grape-Nuts product, given that it would no longer fear losing sales and profits to the newly acquired Nabisco Shredded Wheat product. In a lengthy opinion, the court examined issues of market definition, market shares, traditional coordinated or collusive effects of the acquisition, and the alleged unilateral effects of the transaction. As a factual matter, the court found that Grape-Nuts and Shredded Wheat were not sufficiently close substitutes, and that each faced competition from an array of other products. *See* 926 F. Supp. at 352-53. The court credited the defendant's economic expert Professor Daniel Rubenfeld, who testified that an increase in the price of Grape-Nuts would lead to only a very small increase in the sales of Nabisco Shredded Wheat cereals. *See* 926 F. Supp. at 356. The court concluded that "it would not be profitable for Kraft to raise the price of Grape-Nuts in the expectation that a substantial portion of its lost sales would go to Nabisco Shredded Wheat, because it is likely that the lost sales would be dispersed among a wide variety of products, and that Nabisco Shredded Wheat would gain only a small percentage of those losses. The State has failed to prove its claim of adverse unilateral effects." 926 F. Supp. at 366.

In *Gillette*, the government sought a preliminary injunction enjoining Gillette's proposed acquisition of Parker. The government challenged the deal because of alleged adverse unilateral effects in the premium fountain pen market — refillable fountain pens with a retail price between $50 and $400. The district court rejected the argument. "[T]here is ample evidence that the merged company will not be able to increase prices on premium fountain pens unilaterally." 828 F. Supp. at 84. According to the court, fountain pens compete with other modes of writing. In essence, the court held that the diversion ration between Gillette's and Parker's premium fountain pens was not high. Finally, the court reasoned that there were no substantial barriers to entry, permitting new companies to enter the market or permitting existing companies to reposition their existing product lines in response to a unilateral price increase.

The fact that the plaintiffs in these cases were not successful on the merits does not invalidate the legitimacy of unilateral effects theory, nor does it call into question the ability of the courts to implement these approaches. While potentially complicated, these issues can be litigated and resolved in much the same manner as other complex problems in antitrust cases.

ing the welfare effects of proposed mergers or other restraints exist and are being employed. Second, whether or not the defense of intramarket second-best tradeoffs is embraced, antitrust lawyers and judges will continue to confront increasingly complex economic theories and econometric techniques. That being said, it is important that the use of theory not outstrip the policy framework and heuristics necessary to filter the economic information and process the statistical output. Antitrust law is not a forum to test the limits of economic theory. The theory must fit in an appropriate doctrinal setting and it must be administrable within the context of litigation, which is predicated upon a system of lay factfinders.

C. *Cost-Benefit Analysis and Balancing Error Types*

The desirability of an affirmative defense for intramarket second-best tradeoffs can be assessed in terms of traditional cost-benefit analysis.[140] The social welfare gains from permitting otherwise prohibited business arrangements under contemporary antitrust doctrine are the primary benefits of allowing the defense. Social welfare losses from mistakenly approving welfare-impairing conduct that would have been prohibited but for the affirmative defense, and private and social litigation expenses, are the primary costs.[141]

---

140. For alternative formulations of cost-benefit rules assessing antitrust standards, see Easterbrook, *supra* note 2, at 16 ("The legal system should be designed to minimize the total costs of (1) anticompetitive practices that escape condemnation; (2) competitive practices that are condemned or deterred; and (3) the system itself."); Frank H. Easterbrook, *Workable Antitrust Policy*, 84 MICH. L. REV. 1696, 1711 (1986) ("Overreaching the limits of adjudication will increase the rate of error. We want to hold to a minimum the sum of the costs of harmful activity wrongly condoned and useful activity wrongfully condemned (or discouraged)." (footnote omitted)); Fisher & Lande, *supra* note 23, at 1670-71 (stressing the need to consider type one errors, type two errors, and type three errors, which are defined as the errors that occur "when compliance with merger policy creates excessive costs to businesses, enforcers, and decisionmakers"). *But see* Michael S. Jacobs, *An Essay on the Normative Foundations of Antitrust Economics*, 74 N.C. L. REV. 219, 257 (1995) (arguing that this type of cost-benefit analysis of antitrust policy is inherently indeterminate).

141. Simplistically, an affirmative defense for intramarket second-best claims should be adopted if

$$W(G_{SB}) - W(\text{TypeOneError}_{SB}) > W(\text{TypeTwoError}_{SB}) + \text{Admin \& Litigation Costs}_{SB}$$

where "G_{SB}" represents the universe of all "legitimate" or "good" second-best claims; "$W(G_{SB})$" equals the welfare benefits of permitting all legitimate second-best claims; "$W(\text{TypeOneError}_{SB})$" represents the welfare benefits lost when courts fail to recognize this subset of legitimate claims; "$W(\text{TypeTwoError}_{SB})$" represents the welfare loss, where courts approve illegitimate second-best claims under the guise of the affirmative defense; and "$\text{Admin \& Litigation Costs}_{SB}$" equals the change in administrative and litigation costs associated with adopting the affirmative defense.

Framing the discussion in terms of a formula and defining specific variables are intended to help conceptualize the concerns relevant to assessing the costs and benefits of a new defense, not to attribute any degree of mathematical certainty to the process. In particular, $W(\text{TypeTwoError}_{SB})$ is somewhat loosely defined. In addition to the illegitimate claims actually litigated and mistakenly approved by the courts (classic type two errors), one must

Market failures are common and can disrupt the functioning of isolated industries or entire markets. Ultimately, the importance of creating a second-best tradeoff defense is as great or as small as the problem of unaddressed market failures that second-best solutions can ameliorate. This includes the benefits of conduct that is encouraged because of the defense, but does not lead to antitrust litigation, as well as conduct that does. We cannot expect all theoretically possible gains to be realized because some courts will not recognize valid defenses and some parties may avoid defensible conduct because they fear mistaken verdicts or litigation costs. At the same time, we must also recognize the costs of mistakenly allowing the defense in situations where the private conduct actually reduces total welfare. These costs should be low initially because of the rigorous conditions that must be met to establish the defense.[142]

Importantly, some number of antitrust cases raising second-best claims will be brought and litigated regardless of whether an intramarket second-best defense is acknowledged. Currently, these cases either ignore particular second-best concerns altogether, or raise them under the guise of arguments that have a recognized doctrinal basis. Examples include restraints evaluated under rule of reason, such as R&D joint ventures and restrictions on advertising,[143] but also include

---

account for the welfare loss associated with illegitimate conduct encouraged by the affirmative defense that is never challenged or litigated. There are no added administrative and litigation costs in these cases, but, by definition, this type of conduct results in a net reduction in total welfare. Similarly, in speaking of the benefits of the affirmative defense, one must count not only those cases where welfare-enhancing conduct is challenged and upheld under the affirmative defense, but also those cases where welfare-enhancing conduct is encouraged by the affirmative defense but never challenged or litigated. These cases are accounted for in $W(G_{SB})$, which incorporates the positive welfare effects of all legitimate second-best claims.

142. The capacity of the courts to administer the affirmative defense and the complexity of the rule that is adopted will determine the frequency and magnitude of mistakes. Substantive and procedural aspects of a defense can influence the likelihood and distribution of errors. The affirmative defense outlined in this Article is intended to create a bias in favor of mistaken denials of the defense (type one errors). Making second-best considerations an affirmative defense assigns the burden of proof to the defendant. Imposing a clear and convincing standard of evidence further ensures that the errors associated with the defense will be mistaken denials. These hurdles and biases will reduce the likely expected benefits of the new rule, but should also decrease the costs of the new regime. The difficulties of successfully mounting the defense should also reduce the number of illegitimate claims brought, for there is little sense in going to the expense of mounting the defense if it is unlikely to prevail. Moreover, given the prospect of antitrust liability, fewer bad transactions will be undertaken in the hope of succeeding with an illegitimate second-best defense. Legitimate claims will also be discouraged, but their greater prospects of success means that this should happen at a lower rate. Arguably, the conditions for successfully mounting the defense are too strict. Once courts, enforcers, and parties become better acquainted with what is needed to establish and attack the defense, it might be appropriate to relax some of the requirements, particularly the burden of establishing all conditions for the defense by clear and convincing evidence.

143. See discussion of R&D joint ventures and information market failures *supra* Section I.B.3.

efforts to masquerade alleged benefits of market power as efficiency benefits under the efficiency defense in merger cases.[144]

The primary effect of recognizing an affirmative defense for these cases may be to shift resources from relatively inappropriate arguments under existing doctrine to relatively more appropriate arguments directly addressing the welfare issues that second-best concerns raise. For example, fewer resources might be devoted justifying a broad market definition or proving the absence of market power in merger cases, and more might be spent proving that the conduct addresses endemic market failures and for this reason will increase total welfare.

Struggles over market definition are a flash point in most antitrust cases. Given the important role of market-based presumptions, these battles are often driven more by litigation strategy than theoretical differences about the correct economic market. Defendants typically want an expanded definition of the market, and plaintiffs and regulators want a restrictive definition. When a party seeks to make an intramarket second-best defense, however, the effect will usually be to confess market power. The result will be to shift attention and resources away from disputes about what the market is, to an examination of the economic consequences of the behavior in question.[145] Of course, new strategic incentives will also arise, potentially leading defendants to argue for overly narrow economic markets. Courts can best navigate these conflicts if they view the process of market definition as an effort to establish the appropriate boundaries for a defensible partial equilibrium analysis.[146]

The most significant effect on administrative and litigation costs will come from those cases that would not have been brought but for the new doctrine. One must account for the full transaction costs associated with these cases. Defendants in antitrust cases bear their own litigation expenses. The amount they are willing to spend on litigation

144. See discussion of hospital merger cases *infra* note 149.

145. Of course, it is possible that defendants could argue both for an expansive market definition and, in the alternative, for an intramarket second-best defense if a more narrowly defined market is adopted. If this happens, there would be no conservation of resources. While plausible, such litigation strategies are not without risk, and are unlikely to represent the dominant approach.

146. When the market is too narrowly defined, intermarket inefficiencies may undercut intramarket gains. Although I have excluded intermarket second-best considerations from the defense, largely on feasibility grounds, where during the struggle over market definition it becomes obvious that in a plausible alternative market the welfare implications are negative on balance, this could be enough to undercut the defendant's intramarket second-best defense, even though defendants have no burden of proving that such inefficiencies do not exist. This response to the intramarket defense should be limited to plausible larger markets so as not to encourage litigation of intermarket issues. Since plaintiffs and enforcers will usually want to limit market size to establish market power and will still have a large incentive to do so, allowing this incidental assessment of larger market inefficiencies when the larger market represents a plausible market definition should not unduly expand litigation.

depends upon the expected profits of the underlying activity and the probability of a successful antitrust outcome. Ideally, defendants would make litigation decisions based on the social rather than the private benefits flowing from their conduct, and ideally, defendants would internalize all rather than just some of the social costs of litigating disputes.[147] By definition, legitimate second-best claims are associated with net increases in social welfare. This is conduct society wants to allow and cases that society wants to have litigated if challenged (assuming that the welfare gain exceeds the social costs of litigation). Unfortunately, private parties will spend no more in litigating a claim than the expected value of the private producer surplus. To the extent that producer surplus differs from the total welfare associated with the conduct, private parties will have incentives either to overlitigate or underlitigate these claims.[148] While there is no guarantee that private parties will invest the socially optimum amount in litigation, we can at least be reassured with legitimate claims that the social and private incentives are aligned in the same direction.

Illegitimate second-best claims raise distinctly different questions. These cases represent a double loss. Litigating these cases consumes private and social resources, and to the extent that courts make mistakes and approve these illegitimate claims, the private conduct will further reduce social welfare. The specter of illegitimate claims justifies structuring the defense initially, as I have argued, in a manner that minimizes the risk of mistakenly allowing the defense. The prospect of mistaken victories invites private parties to engage in conduct and defend cases that should not be defended, and provides incentives to increase the resources invested in litigation. Reducing the risks of mistaken defense victories minimizes both the social costs of the underlying welfare-impairing activity and the administrative costs of litigating these cases.

147. Private parties fail to internalize all the social costs of litigation. Defendants in antitrust cases do not bear the government's costs of investigating and trying claims or the administrative costs of the courts. Some antitrust rules counteract these effects. Damages in antitrust cases are trebled, and defendants must pay the attorney's fees of successful private parties. Moreover, unsuccessful defendants face the prospect of potentially burdensome consent decrees and injunctions that could constrain prospective business conduct. These rules could either undercorrect or overcorrect the problem. Antitrust damages and remedies increase the private cost of antitrust violations and deter undesirable forms of conduct, but given the reality of type one errors, such rules can deter desirable conduct as well.

148. Private parties will underlitigate cases in which the underlying conduct increases both producer and consumer surplus (in these cases, the increase in total welfare will be greater than the increase in producer surplus). Private parties will have incentives to overlitigate cases where the underlying conduct increases total welfare and producer surplus, but decreases consumer surplus (in these cases, the increase in producer surplus will be greater than the net increase in total welfare). If one believed that the positive welfare effects of conduct that increases both producer and consumer surplus (conduct that would be permitted under both Lande's consumer surplus rule and a total welfare standard) were relatively easier for courts to identify, then the danger associated with underlitigating these claims may not be significant.

The actual costs and benefits of adopting an affirmative defense for intramarket second-best tradeoffs are not easy to quantify, and no obvious policy recommendation falls out of this framework. If one believes that the welfare effects of legitimate second-best claims are substantial and that courts can successfully identify such claims, then the case for instituting a defense is relatively strong. If one believes that creating a defense will invite private parties to masquerade anticompetitive behavior as second-best claims and that courts will often not detect such deceptions, then one should have serious reservations about an affirmative defense. Alternatively, if one believes that courts are already struggling with many cases that raise intramarket second-best concerns, but lack an effective framework to assess the real merits of such cases, then an affirmative defense may be a better way than current standards to resolve existing controversies.[149] From this perspective, the costs involved would reflect a better expenditure of existing resources rather than an increase in litigation costs. My contention is that the uncertainty associated with the cost-benefit analysis does not justify inaction, but rather justifies adopting a conservative defense that would produce information and experience to shape future policy assessments.

III. RECONCILING A TOTAL WELFARE STANDARD WITH ANTITRUST JURISPRUDENCE

The question of whether to allow intramarket second-best tradeoffs highlights both the tension between an economically oriented antitrust law and its statutory roots, and the tension between enhancing total welfare (efficiency) and promoting competition as goals of antitrust law enforcement. Among the questions to be confronted are: Can second-best arguments be reconciled with the text and legislative history of the antitrust laws? Are other antitrust law goals inconsistent with embracing a total welfare standard? Must concepts of total welfare and efficiency yield in practice to a more structuralist agenda of protecting a particular vision of "competition"? If the courts implement a total welfare standard, what is the proper balance between static and dynamic efficiency concerns? When should courts demand

149. Hospital mergers are a persuasive example. Hospital mergers are taking place in markets rife with market failures. Recognizing an affirmative defense for second-best claims is unlikely to increase the number of these mergers. A second-best defense, however, would provide a more effective framework in which to assess the merits of these cases. In its absence, courts tend either to mischaracterize the lower costs from reduced nonprice competition as increases in efficiency without appreciating the corresponding reduction in consumer surplus, or to focus only on the adverse effects on consumer surplus and fail to consider the private and social savings from conserving resources. An intramarket second-best defense and an examination of the effects of merger on total welfare would result in better decision-making and a more effective use of litigation resources in these cases. *See* Hammer, *supra* note 23 (providing a critical examination of the hospital merger case law).

first-best solutions to problems in markets with substantial market failures? How will permitting second-best solutions that enhance market power affect the ability of private parties and legislatures to undertake broader reform? Is an inquiry into second-best tradeoffs consistent with the proper institutional role of the courts?

My intention here is not to resolve all of these issues but to suggest that seeking answers to these questions should be the subject of serious antitrust scholarship. Unlike Judge Bork, I do not contend that a total welfare standard is statutorily required,[150] nor that the ratification of a total welfare standard represents the necessary or inevitable culmination of an economically oriented antitrust law. Rather, my contention is that a total welfare antitrust standard constitutes an increasingly "living option" under an organic approach to antitrust law *and* that such a standard can be reconciled with the courts' institutional role. Furthermore, I argue that sound policy rationales support

150. Judge Bork advances the strong claim that, as a matter of statutory interpretation, total welfare is the exclusive goal of the Sherman Act. According to Bork, the relevant evidence establishes "conclusively that the legislative intent underlying the Sherman Act was that courts should be guided exclusively by consumer [total] welfare and the economic criteria which that value premise implies." Bork, *supra* note 21, at 11; *see also* BORK, *supra* note 12, at 50-51, 61-66. While Bork speaks of "consumer welfare," his understanding of this term is the equivalent of what I define in this Article to mean total welfare. *See* discussion *supra* note 21. The statutory basis for a total welfare standard has been almost universally rejected by antitrust scholars. *See, e.g.*, Thomas C. Arthur, *Farewell to the Sea of Doubt: Jettisoning the Constitutional Sherman Act*, 74 CAL. L. REV. 263, 290 n.125 (1986) (criticizing Bork's strong consumer welfare and efficiency claims); John J. Flynn, *The Reagan Administration's Antitrust Policy, "Original Intent" and the Legislative History of the Sherman Act*, 33 ANTITRUST BULL. 259, 267 ("Everyone who has made a considered study of the legislative history of the major antitrust laws flatly rejects Judge Bork's assertion that 'consumer welfare' was the only goal Congress had in mind when it enacted the Sherman Act."); Eleanor M. Fox, *Consumer Beware Chicago*, 84 MICH. L. REV. 1714, 1715 n.5 (1986) (criticizing similar usage of the term "consumer" in references to the legislative history); Herbert Hovenkamp, *Antitrust's Protected Classes*, 88 MICH. L. REV. 1, 22 (1989) ("But Bork's analysis of the legislative history was strained, heavily governed by his own ideological agenda.... Not a single statement in the legislative history comes close to stating the conclusions that Bork drew."); Hovenkamp, *supra* note 12, at 250 ("Bork's work has been called into question by subsequent scholarship showing that in 1890 Congress had no real concept of efficiency and was really concerned with protecting consumers from unfavorable wealth transfers."); Edwin J. Hughes, *The Left Side of Antitrust: What Fairness Means and Why It Matters*, 77 MARQ. L. REV. 265, 273-74 (1994) (criticizing Bork's claim that consumer welfare is the exclusive goal of the Sherman Act); Jacobs, *supra* note 140, at 232-34 (discussing and criticizing Bork's strong consumer welfare claim); Louis Kaplow, *Antitrust, Law & Economics, and the Courts*, LAW & CONTEMP. PROBS., Autumn 1987, 181, 206-12 (discussing and criticizing Bork's strong consumer welfare claim); Alan J. Meese, *Price Theory and Vertical Restraints: A Misunderstood Relation*, 45 UCLA L. REV. 143, 155-58 (1997) (detailing the "populist" critique to Bork's consumer welfare claim); Millon, *supra* note 68, at 1231-35 (careful exposition and critique of Bork's consumer welfare claim); Rudolph J. Peritz, *A Counter-History of Antitrust Law*, 1990 DUKE L.J. 263, 282 n.72 ("Too much has already been expended in demonstrating that Bork's account is mistaken."); *see also* Christopher Grandy, *Original Intent and the Sherman Antitrust Act: A Re-Examination of the Consumer-Welfare Hypothesis*, 53 J. ECON. HIST. 359, 359 (1993) ("Congress appeared to reject consumer welfare. If anything, Congress seemed more concerned with producer, rather than consumer, welfare.").

antitrust law's evolution beyond competition into the realm of express welfare analysis.

A. *Total Welfare: A Living Option in Antitrust Law's Evolution*

Whether a total welfare standard is a living option for antitrust law must be determined with reference to all antitrust statutes — including the Federal Trade Commission Act ("FTC Act") and the Clayton Act — not simply the Sherman Act. The task, however, is not to interpret these laws in isolation, but rather to understand their provisions in the context of the complex relationship that exists between the courts and Congress in their joint oversight of private markets. A relational approach to antitrust law envisions power-sharing between the courts and Congress. The courts have primary responsibility for shaping antitrust doctrine and managing its scope and orientation. Congress retains the authority to intervene at any time, change particular rules, carve out areas of immunity, or impose new standards.

The Sherman Antitrust Act was passed in 1890. Substantively, the Act speaks at a very general level, prohibiting contracts, combinations, and conspiracies in restraint of trade, as well as monopolization and attempted monopolization. In spite of the volumes that have been written about antitrust history, most scholars would concede that it is simply not possible to divine a clear and specific mandate from either the text or the legislative history of the law. Instead, most courts and scholars regard antitrust as a statutorily sanctioned area of federal common law, reflecting an organic rather than a statutory methodological orientation.[151] The text of the statute and the Supreme Court's

151. Most modern scholars, regardless of their substantive antitrust agendas, embrace an organic methodological orientation. *See, e.g.*, William F. Baxter, *Separation of Powers, Prosecutorial Discretion, and the 'Common Law' Nature of Antitrust Law*, 60 TEXAS L. REV. 661, 663 (1982) ("To provide this flexibility, Congress adopted what is in essence enabling legislation that has permitted a common-law refinement of antitrust law through an evolution guided by only the most general statutory directions."); Easterbrook, *supra* note 140, at 1702 ("[The Sherman Act] does not contain a program; it is instead a blank check."); *id.* at 1705 ("The Sherman Act set up a common law system in antitrust."); Kaplow, *supra* note 150, at 181, 213 ("Congress apparently contemplated a common law type of development when it enacted the Sherman Act."); Wiley, *supra* note 29, at 777 ("[The Sherman Act] is most reasonably interpreted to grant courts common law power to initiate substantive policy."). This is also the orientation that has guided the Court for the last seventy years in its interpretation of the Sherman Act. *See, e.g.*, State Oil Co. v. Khan, 522 U.S. 3, 20-21 (1997) ("Thus, the general presumption that legislative changes should be left to Congress has less force with respect to the Sherman Act in light of the accepted view that Congress 'expected the courts to give shape to the statute's broad mandate by drawing on common-law traditions.' " (quoting National Soc'y of Prof'l Eng'rs v. United States, 435 U.S. 679, 688 (1978))); Appalachian Coals, Inc. v. United States, 288 U.S. 344, 359-60 (1933) ("As a charter of freedom, the [Sherman] Act has a generality and adaptability comparable to that found to be desirable in constitutional provisions."). Standing virtually alone in their advocacy of an antitrust doctrine grounded exclusively in statutory interpretation are Robert Bork and Thomas Arthur. *See* Arthur, *supra* note 150, at 267 ("All the contending antitrust schools

precedents establish parameters restricting the law's evolutionary development, but over the past one hundred years, antitrust law has demonstrated tremendous flexibility and has been highly responsive to changes in economic thinking and social policy. Nevertheless, appreciating the history of the various antitrust provisions is important to understanding the law's different possible futures.

The FTC Act and the Clayton Act were adopted in 1914, largely in response to the Supreme Court's 1911 decision in *Standard Oil Co. v. United States*.[152] Advocates of new antitrust legislation criticized the "rule of reason" interpretation of the Sherman Act for being too permissive and for not condemning many forms of objectionable anticompetitive conduct.[153] Some legislators harbored institutional and procedural concerns about the *Standard Oil* decision as well, believing that the "rule of reason" was too vague and created too much discretion for judges.[154] The response, however, was not to repeal or substantively amend the Sherman Act, but to enact new antitrust regulations to supplement its provisions.

The Clayton Act and the FTC Act reflect different legislative strategies. In contrast to the Sherman Act's general prohibition of "restraints of trade," the Clayton Act prohibits specific types of conduct — tying arrangements, exclusive dealing, and certain forms of price discrimination.[155] In addition, section 7 of the Clayton Act condemns mergers where the effect "may be substantially to lessen competition, or tend to create a monopoly." In 1950, the Clayton Act was amended to close a loophole in the original legislation.[156] The 1914

agree on one critical point: that the Sherman Act cannot, and should not, be given a settled meaning derived from traditional statutory sources. They are all wrong."); Thomas C. Arthur, *Workable Antitrust Law: The Statutory Approach to Antitrust*, 62 TUL. L. REV. 1163, 1169 (1988) ("[T]he only way courts can produce stable, rational antitrust law is to replace the constitutional approach with a statutory approach . . . ."). Bork's statutory position is outlined *supra* note 150.

152. 221 U.S. 1 (1911).

153. *See* Lande, *Wealth Transfers*, *supra* note 23, at 126.

154. *See* Arthur, *supra* note 150, at 291 n.132 ("[The rule of reason] substitutes the court in the place of Congress, for whenever the rule is invoked, the court does not administer the law, but makes the law." (quoting from a 1913 Report of the Senate Committee on Interstate Commerce)). Importantly, neither the FTC Act nor the Clayton Act amended the rule of reason announced in *Standard Oil* or expressly cabined the discretion of the judiciary in applying the Sherman Act.

155. *See* HOVENKAMP, *supra* note 12, at 57 (summarizing the scope of the 1914 Clayton Act). The price discrimination provision was amended in 1936 by the Robinson-Patman Act. The Robinson-Patman Act was intended to protect small independent retailers in the face of growing chain stores and represents a value choice by Congress for something other than the strict promotion of economic efficiency. Modern interpretations of the Robinson-Patman Act, however, further illustrate the transformative power of an organically oriented antitrust methodology. *See* discussion *infra* note 174.

156. For a general discussion of the 1950 Celler-Kefauver amendments and the circumstances surrounding their passage, see HOVENKAMP, *supra* note 12, at 58-59; Bok, *supra* note 2, at 226; and Lande, *Wealth Transfers*, *supra* note 23, at 130-42.

provisions applied only to stock acquisitions and could easily be circumvented by arranging transactions as asset acquisitions.[157] Significantly, in addition to ensuring that section 7's prohibition would apply to all mergers regardless of form, the 1950 amendments marked a strong and renewed congressional condemnation of economic concentration more generally.

The 1914 FTC Act, more than the Clayton Act, adopted an approach similar to that of the original Sherman Act. Rejecting proposals for crafting a detailed list of prohibited anticompetitive acts, section 5 of the FTC Act broadly condemns "unfair methods of competition." Power to enforce this provision was vested in a new administrative body that could bring expert knowledge to the problem of evaluating business conduct.[158] Importantly, however, this Commission was to share the function of enforcing the nation's antitrust laws with the Department of Justice, the federal courts, and private parties empowered to sue independently for antitrust violations. The policy remained one of concurrent and redundant systems of enforcement.

For the first time, the FTC Act and the Clayton Act wrought the language of "competition" into the textual fabric of the antitrust laws. Justice Holmes's 1904 observation that antitrust law has nothing to do with competition was no longer valid.[159] Even those embracing an organic approach to antitrust law must pause and consider what it means to engage in "unfair methods of competition" or to "substantially lessen competition." Unfortunately, there is little in the text or the legislative history to answer these questions precisely.[160] Competition in these debates has the same flavor of "rivalry" that runs from Adam Smith to the Sherman Act.[161] Even though formal neoclassical models of competition were well developed by 1950, these more structured

157. *See* Lande, *Wealth Transfers, supra* note 23, at 130.

158. For a general discussion of the 1914 FTC Act and the circumstances surrounding its passage, see HOVENKAMP, *supra* note 12, at 57, and Lande, *Wealth Transfers, supra* note 23, at 107-26.

159. *See* Northern Sec. Co. v. United States, 193 U.S. 197, 403 (1904) (Holmes, J., dissenting).

160. Typical of the vacuous statements in the Record is that of Representative Morgan speaking in defense of the Clayton Act: "[T]he one thing we wish to maintain, and retain and sustain, is competition. We want to destroy monopoly and restore and maintain competition." 51 CONG. REC. 9265 (1914).

161. Classical economists understood competition in terms of "rivalry." *See* George J. Stigler, *Perfect Competition, Historically Contemplated*, 65 J. POL. ECON. 1, 1 (1957) (" 'Competition' entered economics from common discourse, and for long it connoted only the independent rivalry of two or more persons."). This same basic understanding is reflected in the legislative history of the Sherman Act. *See* Hovenkamp, *supra* note 150, at 23 ("Although 'competition' was never defined formally, most of the speakers appeared to use it to mean 'rivalry,' or the presence of multiple sellers in the market. Never once was 'competition' defined as a state of affairs in which price equals marginal cost, or any other measure of cost.").

understandings of competition seem hardly to have influenced the congressional discussions. Indeed, the 1950 debate surrounding competition has more political than economic content.[162] As Derek Bok concluded, "it seems abundantly clear that 'competition' meant far more to Congress than prices, costs, and product innovation."[163] One finds little additional guidance as to the meaning of competition if one turns from the language of the law to the underlying set of legislative policies and objectives.[164]

---

162. *See* Bok, *supra* note 2, at 236 ("To anyone used to the preoccupation of professors and administrators with the economic consequences of monopoly power, the curious aspect of the debates is the paucity of remarks having to do with the effects on concentration of prices, innovation, distribution, and efficiency. To be sure, there were allusions to the need for preserving competition. But competition appeared to possess a strong socio-political connotation which centered on the virtues of the small entrepreneur to an extent seldom duplicated in economic literature."); Hovenkamp, *supra* note 12, at 254.

163. Bok, *supra* note 2, at 248; *see also* Brown Shoe Co. v. United States, 370 U.S. 294, 315 (1962) ("The dominant theme pervading congressional consideration of the 1950 amendments was a fear of what was considered to be a rising tide of economic concentration in the American economy."); Bok, *supra* note 2, at 235-36 ("This situation was appraised in the same Jeffersonian, egalitarian fashion by almost all who spoke for the bill.... In the minds of the Congressmen, the growth of these large economic groups could lead only to increasing government control; freedom would corrode and the nation would drift into some form of totalitarianism." (footnote omitted)); Kaplow, *supra* note 150, at 211 ("[T]he 1950 amendment to section 7 of the Clayton Act[] was expressly directed at social and political aspects of economic concentration, rather than enhancing economic efficiency." (footnote omitted)). Similar statements can be found in the debates leading up to the passage of the original Clayton Act. *See* James May, *Antitrust in the Formative Era: Political and Economic Theory in Constitutional and Antitrust Analysis, 1880-1918*, 50 OHIO ST. L.J. 257, 297 (1989) ("Fears of political domination by big businesses were powerfully reiterated in Congress during the 1914 Clayton Act debates, both in report language and in comments by individual congressmen."); *id.* ("Senator Borah's dire 1914 warning that monopolies 'divide our people into classes, breed discontent and hatred, and in the end riot, bloodshed, and French revolutions.' " (quoting 51 CONG. REC. 15,955 (1914) (statement of Senator Borah))).

164. No single, consistent purpose emerges. A number of themes are present, some complementary, others conflicting. Robert Lande catalogues the objectives of the FTC Act to include improving economic efficiency, *see* Lande, *Wealth Transfers, supra* note 23, at 108-12, protecting consumers from unfair transfers of wealth, *see id.* at 112-14, preventing unethical business practices, *see id.* at 115-18, curbing the social and political power of monopolists, *see id.* at 118-19, protecting small businesses, *see id.* at 120-21, and providing consumer protection, *see id.* at 121-26. Lande lists the goals of the Clayton Act to include improving economic efficiency, *see id.* at 131-35, protecting consumers from unfair wealth transfers, *see id.* at 135-36, preventing economic concentration, *see id.* at 137-38, and protecting small businesses, *see id.* at 139-40. Lande's contention that efficiency concerns were acknowledged in the 1950 legislative debates is in tension with Bok's conclusion that "[t]here is little basis for concluding that the achievement of lower costs as such should give rise to favorable treatment under section 7. The possibility of lower costs was brushed aside in the legislative deliberations and there is every reason to believe that Congress preferred the noneconomic advantages of deconcentrated markets to limited reductions in the cost of operations." Bok, *supra* note 2, at 318 (footnotes omitted). Even Lande acknowledges that there was a striking lack of appreciation for situations in which sentiments in favor of lower prices or against economic concentration may conflict with efficiency concerns. *See* Lande, *Wealth Transfers, supra* note 23, at 135 n.266 ("There was apparently no explicit analysis of the possible tradeoffs involved in implementing actions that simultaneously increase efficiency and raise prices to consumers."). For a further assessment of the legislative history of the Clayton Act that concludes that productive efficiency considerations are relevant, see

Ratification of an affirmative defense for intramarket second-best tradeoffs ultimately requires the objective of total welfare to trump structural understandings of competition. Can an organically oriented antitrust law legitimately pursue total welfare at the occasional expense of "competition"?[165] While such an outcome is not dictated by the text or the legislative history of the antitrust laws, is it precluded? The fact that there is no rigid or even clear vision of "competition" embedded in the text of the antitrust laws opens the door to this possibility. More importantly, the courts have consistently interpreted the various antitrust statutes in a manner that has avoided rigid structural rules, has preserved flexibility, and has afforded the judiciary substantial discretion in its administration. Section 1 of the Sherman Act condemns only "unreasonable" restraints of trade, acknowledging that the essence of every contract is to restrain trade in some sense.[166] Section 2 of the Sherman Act prohibits monopolization or attempted monopolization, acknowledging that some monopolies can be acquired and perpetuated that would not be illegal.[167] Section 5 of the FTC Act prohibits "unfair methods of competition," but rather than triggering a new set of antitrust rules or a unique jurisprudence centered on the concept of "competition," section 5 largely defines legality and illegality in terms of the conduct that is permitted or prohibited under the Sherman Act. "Competition," as such, has little independent significance.

Timothy J. Muris, *The Efficiency Defense Under Section 7 of the Clayton Act,* 30 CASE W. RES. L. REV. 381, 393-402 (1980).

165. The tension between competition and efficiency is essentially about whether legal presumptions based on structural understandings of competition are rebuttable. If efficiency is the ultimate objective of antitrust law and competition is simply a proxy for determining whether conduct is efficient, then the presumption that conduct is inefficient because it is anticompetitive should be rebuttable by direct evidence that the conduct in question is welfare-enhancing. *See* Hammer, *supra* note 23 (examining the ability to rebut theoretical antitrust presumptions with the introduction of direct economic testimony). If competition rather than efficiency is the ultimate goal of antitrust law, then the presumption of illegality predicated on the fact that conduct is inconsistent with the court's structural understanding of competition should be irrebuttable. Under an efficiency orientation, the courts would be able to entertain an affirmative defense predicated on intramarket second-best tradeoffs. Under a competition orientation, such a defense may not be available.

166. There is some overstatement in this characterization of section 1. Section 1's per se rules are supposed to be the archetype of rigid categorical distinctions. Even in these rigid categories, one can observe flexibility over time, as the hard per se rules against tying arrangements and group boycotts have eroded. More importantly, there is substantial flexibility in characterizing behavior and therefore in deciding whether to invoke the per se condemnation. *See* discussion *infra* note 194.

167. *See, e.g.,* United States v. Grinnell Corp., 384 U.S. 563, 570-71 (1966); United States v. E.I. du Pont de Nemours & Co., 351 U.S. 377, 389-91 (1956); United States v. ALCOA, 148 F.2d 416, 429 (2d Cir. 1945). In addition, "monopoly conduct" is increasingly being defined in terms of whether there is a legitimate business justification for the action being challenged. *See* Aspen Skiing Co. v. Aspen Highland Skiing Corp., 472 U.S. 585, 608 (1985).

The most difficult case is presented by section 7 of the Clayton Act, which prohibits mergers that "substantially lessen competition or tend to create a monopoly." While the Clayton Act fails to define "competition," scholars like Derek Bok persuasively argue that the law implicitly adopted a structural understanding of competition similar to the economic model of pure competition with its assumed large number of competitors. As Bok points out, "[u]nderlying the legislative deliberations was the conviction that small business and the dispersion of economic power are salutary and should be encouraged by the new section 7. This premise evokes a structural theory of competition which stresses the advantages of larger numbers of small-sized firms."[168]

Thus, the fairest conclusion is that in 1950, Congress, motivated primarily out of basic fears of economic concentration, wanted an antitrust merger policy that prohibited large acquisitions and that preserved opportunity for small independent competitors. Initial judicial enforcement of the Clayton Act was premised upon a structural un-

168. Bok, *supra* note 2, at 247. Hovenkamp has reached a similar conclusion. " 'Competition' within the meaning of the statute does not refer to a state of affairs in which prices are driven to marginal cost and firms are encouraged to pursue all economies in production and distribution. Rather it refers to a regime in which small businesses have a chance to compete against larger, more efficient rivals." Hovenkamp, *supra* note 12, at 254 (basing his argument on a reading of the legislative history of the 1950 Cellar-Kefauver amendments). Not surprisingly, Robert Bork provides a strikingly different definition of what "competition" means. After reviewing possible definitions of competition in terms of rivalry, the absence of restraint, the economist's conception of perfect competition, and the maintenance of a fragmented industry structure, *see* BORK, *supra* note 12, at 58-60, Bork defines competition "as a shorthand expression, a term of art, designating any state of affairs in which consumer [total] welfare cannot be increased by moving to an alternative state of affairs through judicial decree," *id.* at 61. Bork operationally defines competition as the maximization of total welfare. In essence, Bork resolves the potential conflict between structural understandings of competition and economic efficiency by abandoning a structural approach and defining competition in terms of total welfare (economic efficiency). This tactic has been criticized for failing to acknowledge competition as an intrinsic value. *See* Peritz, *supra* note 150, at 303-04 ("In short, the traditional view of competition as an important end in itself is turned on its head: Competition is valued only when it serves wealth maximization. That is, competition is valued only as a means for increasing the cumulative market value of private property."); *see also* David W. Barnes, *Nonefficiency Goals in the Antitrust Law of Mergers*, 30 WM. & MARY L. REV. 787, 849-53 (1989) (criticizing Bork's definitions of competition and consumer welfare for being too narrowly focused). Others have criticized this approach for failing to be faithful to the legislative history of the various antitrust laws and even the common parlance of professional economists.

> Economic theorists go about proving that competition produces efficiency in some circumstances and inefficiency in others; the language by which they describe their efforts clearly reveals that they use "competition" to describe the process of interaction (existence of rivalry, specified in various ways) and "efficiency" to characterize the properties of the result of many processes, of which competition is only one.

Kaplow, *supra* note 150, at 210. Kaplow argues that the "efficiency-only interpretation also is clearly contrary to the legislative history of all the antitrust statutes." *Id.*

derstanding of competition that supported a series of presumptions based upon levels of economic concentration.[169]

Section 7 represents the high-water mark for a legislatively asserted antitrust policy predicated upon a vision of enhancing competition. But its application over the years illustrates the evolutionary potential of antitrust law to transform its own initial understandings. The evolution of section 7 jurisprudence reveals a steady transition away from structural understandings of competition and towards a total welfare orientation. The presumptions triggered by market share data can be rebutted by an increasing number of factors.[170] Proof of low barriers to entry, for example, can rebut a presumption of illegality, undermining the assertion that "competition" only exists when there is a large number of active businesses.[171] Additionally, the evolution of the "efficiency" defense under section 7 in the lower courts reveals that structural understandings of competition are increasingly being subordinated to broader understandings of social welfare.[172]

---

169. *See, e.g.*, United States v. Philadelphia Nat'l Bank, 374 U.S. 321, 363 (1963) ("[A] merger which produces a firm controlling an undue percentage share of the relevant market, and results in a significant increase in the concentration of firms in that market, is so inherently likely to lessen competition substantially that it must be enjoined in the absence of evidence clearly showing that the merger is not likely to have such anti-competitive effects.").

170. *See* Hammer, *supra* note 23 (discussing recent cases that assert a wide range of factors as being capable of rebutting the prima facie case of illegality based on market share presumptions, including 1) ease of entry, 2) efficiency, 3) flailing company or weak competitor status, 4) the presence of large sophisticated buyers, 5) nonprofit status, 6) changing market conditions, 7) direct evidence that mergers will not increase prices, and 8) promises by the merging parties not to behave anticompetitively).

171. *See, e.g.*, FTC v. Cardinal Health, Inc., 12 F. Supp. 2d 34, 55 (D.D.C. 1998) ("A court's finding that there exists ease of entry into the relevant product market can be sufficient to offset the government's prima facie case of anticompetitiveness."); United States v. Long Island Jewish Med. Ctr., 983 F. Supp. 121, 149 (E.D.N.Y. 1997) ("A merger is not likely to cause an anti-competitive effect if other participants can enter the relevant markets and reduce the likelihood of a price increase above competitive levels."); United States v. United Tote, Inc., 768 F. Supp. 1064, 1071 (D. Del. 1991) ("United Tote's second argument is that it is so easy to enter the totalistor market that high market share does not accurately reflect an ability to exercise market power."); McCaw Personal Communications, Inc. v. Pacific Telesis Group, 645 F. Supp. 1166, 1174 (N.D. Cal. 1986) ("The existence of low barriers to entry may rebut a prima facie showing of illegality, even when the combined market shares of the merged firms is quite high.").

172. *See, e.g.*, FTC v. University Health, Inc., 938 F.2d 1206, 1222 (11th Cir. 1991) ("We conclude that in certain circumstances, a defendant may rebut the government's prima facie case with evidence showing that the intended merger would create significant efficiencies in the relevant market."); *Long Island Jewish Med. Ctr.*, 983 F. Supp. at 146-47 (entertaining hospital's argument that efficiencies could rebut the presumption of anticompetitive effects); FTC v. Butterworth Health Corp., 946 F. Supp. 1285, 1300-01 (W.D. Mich. 1996) (efficiency gains can rebut the presumption of anticompetitive effects), *aff'd without op.*, 121 F.3d 708, *reported in full*, 1997-2 Trade Cas. (CCH) ¶ 71,863 (6th Cir. 1997); United States v. Country Lake Foods, Inc., 754 F. Supp. 669, 680 (D. Minn. 1990) ("The Court finds these efficiencies relevant, not so much as an independent factor justifying the proposed acquisition, but as further evidence that the proposed acquisition will enhance competition."). *But see* FTC v. Staples, Inc., 970 F. Supp. 1066, 1089 (D.D.C. 1997) ("Assuming that it is a viable defense, however, the Court cannot find in this case that the defendants' efficiencies evidence rebuts the presumption that the merger may substantially lessen competition or shows that the

Is the evolution of the courts' section 7 jurisprudence legitimate?[173] This depends, in part, upon one's methodological orientation. It is difficult to reconcile the evolution of the efficiency defense with the text and legislative history of the Clayton Act and the 1950 Celler-Kefauver amendments, just as it would be difficult to reconcile a total welfare standard or intramarket second-best tradeoffs with a strict statutory orientation. Regardless of the ambiguity in the text and lack of a precise definition of "competition," the total welfare standard I presented in Part I does not appear to be a "living option" under a strict statutory approach that accepts a structuralist view of section 7. The fiction that the increases in efficiency and total welfare are "procompetitive," while alluring, is difficult to reconcile with either economic or legislative understandings of "competition." But, as I have just argued, antitrust jurisprudence, influenced by economic theory, has moved beyond this. The courts have treated the antitrust laws as organic in the sense that decisions allow for the informing of antitrust law by economic theory so as to "grow" a body of law that is at times increasingly independent of its original understandings.[174] Under an organic methodological orientation, the evolution of the efficiency defense in the lower courts can claim legitimacy. Here, it is sufficient that the courts' interpretation is not prohibited by the statute and that the doctrine is sound as a matter of antitrust policy, judicial admini-

---

Commission's evidence gives an inaccurate prediction of the proposed acquisition's probable effect.").

173. Organic approaches to antitrust must openly struggle with the question of legitimacy of proposed doctrinal changes. Thomas Arthur has stressed this point. *See* Arthur, *supra* note 151, at 1168 ("Thus, although the constitutional approach to antitrust raises the same concerns of institutional competence and constitutional legitimacy, there is no counterpart in antitrust to the long running tempest in constitutional law over the legitimacy of wholesale judicial lawmaking — even though it is a natural topic for inclusion in the continuing debates over antitrust fundamentals." (footnote omitted)). While I agree with Arthur that questions of legitimacy are important, I do not agree that antitrust legitimacy must be statutorily grounded. Within an organic orientation, antitrust legitimacy is obtained by crafting functional policies that address modern business needs in a fashion that is consistent with contemporary understandings of the institutional role and capacity of the courts.

174. The fate of the Robinson-Patman Act is illustrative. As originally enacted, the anti-price discrimination provisions of the of the Robinson-Patman Act reflected a policy of protecting small independent retailers from encroaching chain stores. As antitrust doctrine has evolved under an efficiency mantel, the desire to safeguard competition and enhance total welfare has largely trumped countervailing desires to protect individual competitors or small businesses. While judges cannot overturn legislation they disagree with, the courts have systematically narrowed the reach of the Robinson-Patman Act. Moreover, in an effort to reconcile the objectives of the Act with the efficiency-oriented provisions of other antitrust laws, courts have clearly indicated that they will not let the Robinson-Patman Act interfere with the procompetitive mandates of antitrust law, even if this creates nearly irreconcilable dilemmas for businesses. *See, e.g.*, United States v. United States Gypsum Co., 438 U.S. 422, 458 (1978). In a similar move to limit the effects of the Act, the enforcement agencies now devote minimal resources to enforcing its provisions. For a further discussion of the history and development of the Robinson-Patman Act, see HERBERT HOVENKAMP, FEDERAL ANTITRUST POLICY: THE LAW OF COMPETITION AND ITS PRACTICE § 14.6, at 571-74 (2d ed. 1999).

stration, and the evolution of antitrust doctrine. Under this approach, a total welfare standard and a role for intramarket second-best trade-offs remain "living options." Whether such potential should be realized ultimately depends upon whether an affirmative defense would constitute sound policy and whether it could be operationalized in a manner that can be reconciled with the courts' broader institutional role.

B. *Total Welfare: Partnership, Legitimacy, and the Institutional Role of the Courts*

As stressed earlier, American antitrust law embodies a complex partnership between Congress and the courts. What is striking if one looks back over the past one hundred years is not that the FTC Act and Clayton Act were adopted in 1914 and the Clayton Act amended in 1950, but rather that there have been no successful efforts to repeal or enact major substantive changes to the core provisions of the Sherman Act. The dominant legislative trends have been to expand the scope of antitrust liability, to opt for concurrent systems of enforcement, and not to interfere with the organic evolution of antitrust doctrine or the discretion enjoyed by courts in its enforcement.

This does not mean that Congress has been an inactive or silent antitrust partner. Antitrust law represents a default set of rules governing private markets. By express or implied repeal, Congress can displace the antitrust regime in its entirety, with respect to either discrete activities or whole industries. Rather than amending or altering the default antitrust rules, Congress is more likely to supplant antitrust law in those areas it decides to regulate. Thus a common legislative response has been to enact limited exceptions to the antitrust laws, such as the Health Care Quality Improvement Act of 1986, the Local Government Antitrust Act of 1984, the McCarran-Ferguson Act exempting the business of insurance, the National Cooperative Research and Production Act of 1993, the Norris-LaGuardia Act for labor, and the Newspaper Preservation Act.

Congress is not limited to speaking through the text of its statutes. In deciding to overrule cases or chart new directions, the Court is sensitive to a wide variety of legislative signals.[175] The per se rule against

175. *See, e.g.*, Continental T.V., Inc. v. GTE Sylvania, Inc., 433 U.S. 36, 51 n.18 (1977) ("Furthermore, Congress recently has expressed its approval of a *per se* analysis of vertical price restrictions by repealing those provisions of the Miller-Tydings and McGuire Acts allowing fair-trade pricing at the option of the individual States. Consumer Goods Pricing Act of 1975, 89 Stat. 801, amending 15 U.S.C. § 1, 45(a). No similar expression of congressional intent exists for nonprice restrictions."); State Oil Co. v. Khan, 522 U.S. 3, 19 (1997) ("In the context of this case, we infer little meaning from the fact that Congress has not reacted legislatively to *Albrecht*. In any event, the history of various legislative proposals regarding price fixing seems neither clearly to support nor to denounce the *per se* rule of *Albrecht*.").

minimum resale price maintenance has not yet been overturned, largely due to perceived congressional support for the standard. Congress can influence antitrust doctrine by the type of legislation it introduces, the hearings it holds, and the riders that it attaches to appropriations bills.[176] Moreover, antitrust enforcement priorities, policies, and guidelines are politically influenced by the objectives of the sitting President, as well as by the likelihood that controversial actions might provoke new legislation.[177]

On their side of the partnership, the courts have embraced antitrust as a statutorily sanctioned area of common law development,[178] although there are important differences between antitrust law and traditional common law disciplines. One is antitrust's statutory origins. Ultimately, courts are constrained by the antitrust statutes, but the ambiguities wrought into the law and the level of generality at which the antitrust provisions speak mean that this constraint is a loose one. A second difference is the type of sources that courts look to in their decisionmaking. Academic scholarship[179] and economic

176. In the 1980s, Congress passed a number of appropriation riders preventing the Reagan administration's Department of Justice from advocating overruling the per se rule against resale price maintenance announced in *Dr. Miles Medical Co. v. John D. Park & Sons Co.*, 220 U.S. 373 (1911). *See* Act of Nov. 28, 1983, Pub. L. No. 98-166, § 510, 97 Stat. 1102-03; Act of Oct. 18, 1986, Pub. L. No. 99-500, § 605, 100 Stat. 1783, 1783-73; Act of Sept. 30, 1988, Pub. L. No. 100-202, § 605, 101 Stat. 1329-1, 1329-38. In 1985, Congress passed a "sense of the Congress" resolution reaffirming the validity of the per se rule. *See* Act of Dec. 13, 1985, Pub. L. No. 99-180, § 605, 99 Stat. 1169-70. Today, *Dr. Miles,* as it applies to minimum resale price maintenance, is still on the books.

177. The change in enforcement agency policy in the 1994 and 1996 health care statements, from a strict per se treatment of joint price setting by physician networks in the absence of financial risk sharing to a more liberal rule permitting joint pricing in the presence of substantial clinical integration, was arguably influenced by fears of legislative activity. For a discussion of the rule change, see Hammer, *supra* note 45. Some scholars had sharply criticized the 1994 statements for, among other things, the possibility that they would invite unfavorable congressional reform. *See* Clark C. Havighurst, *Are the Antitrust Agencies Overregulating Physician Networks?*, 8 LOY. CONSUMER L. REP. 78, 92-94 (1996); *id.* at 93 ("Unfortunately, unwise administration of the antitrust laws, either by the agencies or by the courts, invites Congress to intervene on behalf of politically powerful physician interests and to enact confusing, possibly over broad correctives or destructive immunities . . . .").

178. *See Khan*, 522 U.S. at 20-21; *see also supra* note 151 and accompanying text (examining the common law orientation of antitrust law).

179. In *Continental T.V., Inc. v. GTE Sylvania, Inc.*, 433 U.S. 36 (1977), the Court overruled the per se rule against vertical nonprice restraints that it had announced in *United States v. Arnold, Schwinn & Co.*, 388 U.S. 365 (1967). A significant factor in the Court's decision was the amount of scholarly criticism the *Schwinn* rule had encountered. "Since its announcement, *Schwinn* has been the subject of continuing controversy and confusion, both in the scholarly journals and in the federal courts. The great weight of scholarly opinion has been critical of the decision . . . ." *GTE Sylvania*, 433 U.S. at 47-48. Similar considerations were cited by the Court in *Khan* when it overruled the per se rule against maximum vertical price restraints announced in *Albrecht v. Herald Co.*, 390 U.S. 145 (1968). "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." *Khan*, 522 U.S. at 15. "Just as *Schwinn was* [sic] 'the subject of continuing controversy and confusion' under the 'great weight' of scholarly criticism . . . *Albrecht* has been widely criticized since its

theory[180] play a greater role in the formation of modern antitrust doctrine than in other common law areas. Similarly, the Court has consciously depreciated the weight of its own precedents and the role of stare decisis in antitrust cases. "But '*[s]tare decisis* is not an inexorable command.' In the area of antitrust law, there is a competing interest, well represented in this Court's decisions, in recognizing and adapting to changed circumstances and the lessons of accumulated experience."[181]

Whether the courts should adopt a total welfare standard in antitrust cases must be examined within the context of this complex judicial and legislative partnership. When viewed in relational terms, as opposed to a static question of statutory interpretation, the very role of legislation begins to change. When Congress has acted, it has done so to supplement rather than displace the Sherman Act. Given the organic nature of antitrust's common law orientation, interventions such as the FTC Act, the Clayton Act, and the Celler-Kefauver amendments can be viewed as legislative efforts to affect the trajectory of the law's development. The lapses of time since the passage of these Acts, as well as court decisions and executive and congressional action and nonaction, establish new parameters on the legitimate paths of the law's prospective growth. Indeed, some scholars have suggested an estoppel-type theory of statutory interpretation should be applied to the antitrust laws.[182] From this perspective, the persuasive power of specific legislative enactments can decay over time, and the freedom of courts to experiment with new doctrines in response to contemporary needs can increase, while the legislature is continually free to reintervene.

---

inception." *Khan*, 522 U.S. at 21 (citation omitted) (quoting *GTE Sylvania*, 433 U.S. at 47-48).

180. In *GTE Sylvania*, the Court indicated that economic notions of competition would have to take primacy over political or sociological constructs of competition. "Competitive economies have social and political as well as economic advantages, but an antitrust policy divorced from market considerations would lack any objective benchmarks." *GTE Sylvania*, 433 U.S. at 54 (citing Northern Pac. Ry. Co. v. United States, 356 U.S. 1, 4 (1958)). The Court examined the economic rationality of a variety of vertical nonprice restraints and found them to be meritorious. "[T]here is substantial scholarly and judicial authority supporting their economic utility. There is relatively little authority to the contrary." *GTE Sylvania*, 433 U.S. at 57-58 (footnote omitted). The Court conducted a similar economic analysis in *Khan*. "After reconsidering *Albrecht*'s rationale and the substantial criticism the decision has received, however, we conclude that there is insufficient economic justification for *per se* invalidation of vertical maximum price fixing." *Khan*, 522 U.S. at 18. Interestingly, the Court made this determination on a review of the case law and the literature, and without the solicitation of expert economic testimony at the district court level. *See Khan*, 522 U.S. at 19.

181. *Khan*, 522 U.S. at 20 (quoting Payne v. Tennessee, 501 U.S. 808, 828 (1991) (alteration in original)).

182. *See* Wiley, *supra* note 29, at 777 n.299 ("Congress in turn has cooperated with this allocation of policy responsibility. It has accepted the bulk of the Court's antitrust invention.").

The treatment of economic efficiency in merger cases from the 1950s to the present is consistent with this understanding. The 1950 Celler-Kefauver amendments reflected a strong congressional concern with economic concentration and an implicit hostility to claims of productive efficiency — at least to the extent that increased efficiency was correlated with increased size. The passage of these amendments was a substantial development in antitrust history and represented a concerted legislative effort to affect the trajectory of antitrust's evolutionary growth. The Court was responsive and gave substantial weight to the law's underlying deconcentration policies in shaping antitrust doctrine in the 1960s (inside and outside the merger context). Legality under section 7 was defined in rigid structural terms. Mergers were condemned at their "incipiency" at relatively low levels of economic concentration, and the Court responded to alleged efficiency claims in a hostile manner.[183]

Much has changed since then. Merger analysis has become far less rigid as economic considerations have increased the sophistication (and complexity) of the analysis.[184] The thresholds at which anticompetitive effects are assumed have increased. The Court has embraced economic efficiency as a positive virtue in other areas of antitrust law, and the lower courts are increasingly recognizing efficiency arguments as potential defenses in Clayton Act section 7 cases.[185] Rather than viewing these developments as a repudiation of the Celler-Kefauver amendments, which they might well have been had they occurred in the immediate wake of the law's passage, these developments can be seen as rational adaptations to changing circumstances, including policy signals sent by the legislature. These changes obtain their legitimacy through the persuasiveness of their underlying justifications, the length of time they endure, and the acquiescence of the legislative branch. Moreover, it can be argued that the degree of common law freedom the judiciary possesses to adopt new antitrust approaches to old questions increases with the novelty of new demands, improvements in economic and judicial understandings of the consequences of merger, and the rate at which contrary legislative interventions decay. It is within this environment that the issue of intramarket second-best

183. *See* United States v. Von's Grocery Co., 384 U.S. 270, 277 (1966) ("Congress sought to preserve competition among many small businesses by arresting a trend toward concentration in its incipiency before that trend developed to the point that a market was left in the grip of a few big companies."); Brown Shoe Co. v. United States, 370 U.S. 294, 346 (1962) ("We cannot avoid the mandate of Congress that tendencies toward concentration in industry are to be curbed in their incipiency . . . ."). The merging grocery stores in *Von's Grocery* had a combined 7.5% share of the L.A. market. *See Von's Grocery,* 384 U.S. at 272.

184. See, for example, *DOJ and FTC Merger Guidelines, supra* note 130, for a description of contemporary enforcement agency approaches to assessing the legality of horizontal mergers.

185. *See* cases cited *supra* note 172.

tradeoffs and the broader question of the welfare orientation of antitrust law need to be assessed. Given antitrust developments in the past twenty-five years, I believe that an affirmative defense recognizing intramarket second-best concerns is within the scope of acceptable evolutionary paths of antitrust law and is thus a living option for judges and antitrust policymakers.

Antitrust doctrine, however, must be concerned not only with the absolute capacity of the courts but also with the relative strengths and weaknesses of private parties, courts, Congress, and the enforcement agencies to address the problem of market failures. Many antitrust rules serve implicit channeling functions, dictated as much by institutional needs as by legislative mandate. Antitrust rules help define a division of labor between private markets and public institutions, between Congress and the judiciary, and between administrative regulation and the courts. In this process, Congress and the courts have formed a partnership with antitrust law establishing the default rules that govern private markets. In the absence of express federal or state regulation, these rules are administered at the discretion of the judicial system.

The challenge, therefore, is to articulate a vision of intramarket second-best tradeoffs and a total welfare standard that can be operationalized within this partnership and that can create a workable division of labor between the legislative, enforcement, and judicial functions. Under a total welfare standard, courts are given the task of maximizing welfare on a market-by-market basis. This is a defensible default rule for policing private markets. A total welfare standard does not give the courts unlimited discretion. By adopting an economic orientation and by limiting analysis to intramarket concerns, two important tasks are channeled out of the courts to the legislative body. The first task is establishing priorities for extra-economic objectives and making tradeoffs between economic and extra-economic goals. The second task is making tradeoffs between distinct economic markets. Conceptually, these functions are better suited to the fact-finding and deliberative capacity of a legislative body.

Antitrust rules must also create an appropriate division of labor between spheres of public and private activity. The affirmative defense for intramarket second-best tradeoffs creates a forum in which welfare-enhancing solutions to market failures can emerge from either the public or the private realm. Current antitrust doctrine tends to discourage private initiatives that could ameliorate the consequences of market failures, biasing potential solutions in a public direction. For example, hospitals may seek to merge so that they do not have to spend money wastefully on amenities such as private parking spaces for physicians, but the law may not allow them to effectuate such savings by concentration. Adopting a total welfare antitrust standard and acknowledging intramarket second-best tradeoffs open the door for

private initiatives. Since antitrust law only establishes a set of default rules, state and federal regulators are free to displace private initiatives.[186] Moreover, the focus of the third element of the affirmative defense on dynamic efficiency is designed to approve private initiatives in a manner that will not impair or substantially increase the cost of subsequent public or private initiatives. The result is a system that encourages innovation and maintains flexibility by permitting a wide range of responses to market failures.

Intramarket second-best tradeoffs address market failures occurring in a single economic market.[187] The internal logic of second-best theory, however, naturally extends beyond intramarket tradeoffs to embrace potential intermarket considerations as well. Is it appropriate to limit antitrust law to partial equilibrium analyses (focusing on a single well-defined market), even as the antitrust mandate is expanded to consider tradeoffs between market power and other failures in the same market? Maintaining a judicial focus on intramarket concerns can be justified on both practical and theoretical grounds. To begin with, the intra- and intermarket distinction is not exogenously determined. Market definition is the first step in most antitrust cases. Proper market definition seeks to include the full range of products that are acceptable consumer substitutes, as well as an appropriate range of supply substitutes. In many instances the boundaries established in this process can capture interrelated market failures.[188] Mar-

186. State or federal legislatures may intervene because they believe a public solution would be more efficient, or because they favor a different social outcome. The prospect of polluters or cigarette manufacturers reaping monopoly profits may strike many as offensive and inappropriate. These outcomes can be avoided by legislation without having to condemn second-best tradeoffs more generally.

187. Of course, this distinction assumes that it is meaningful to speak of "markets" as discrete, definable entities. In theory, as well as in practice, the boundaries separating markets are often porous, and general second-best theory highlights the frequently artificial nature of antitrust market definition. This tension is not dissimilar to the considerations an economist confronts in deciding whether to employ a partial or a general equilibrium framework to model economic phenomena, or in defining the appropriate scope of partial equilibrium analysis.

188. The discussion *supra* note 146 was concerned about private parties manipulating the market definition process. Private parties might have incentives to define markets too narrowly to exploit an intramarket second-best defense that may in fact lead to a net welfare loss in a more properly defined market. The concern here is slightly different. It is possible for intermarket second-best problems to undercut the asserted efficiency of intramarket second-best solutions. As a matter of good policy and good economics, courts should strive to use the market definition process as a tool to capture strong economic interrelationships. If market definition is used to correctly frame the economic inquiry, then the claim that maximizing efficiency within such a partial equilibrium context will lead to an increase in total welfare is economically defensible. *See* Williamson, *supra* note 17, at 987 (arguing that the general theory of second best need not undermine the viability of a partial equilibrium antitrust analysis, so long as the partial equilibrium framework captures strong interconnection); *see also* William J. Baumol, *Informed Judgment, Rigorous Theory and Public Policy*, 32 S. ECON. J. 137, 143-45 (1965) (discussing the ability to partition economic inquiries in a manner that captures dominant second-best considerations); discussion *supra* notes 15-18 and accompanying text.

ket definition aside, private economic activity often embodies, is shaped by, or reflects the existence of market failures in a manner that ensures that economic interconnections are brought before the court as a package. For example, externalities or other market failures may shape firm contracting and integration practices or be the subject of trade association rules and policies.

Given the partnership between Congress and the courts, the intra/intermarket distinction need not be perfect to be functional. Every antitrust theory must struggle with defining the proper institutional role of the courts and the appropriate division of labor between Congress, the enforcement agencies, and the judiciary. The task of developing policies and strategies for policing intermarket second-best concerns can properly be relegated to Congress, both because the analysis strains the boundaries of what can be litigated in court,[189] and because intermarket tradeoffs increasingly involve policy choices that are better suited to the political process (comparisons between apples and oranges held by distinctly different constituencies as opposed to comparisons between apples and apples).[190]

---

189. Antitrust policy must be sensitive to the limitations of judicial capacity. Part II argued that the theoretical and applied economic tools necessary to implement a partial equilibrium economic analysis exist and can be made intuitively accessible to judges and jurors in assessing intramarket second-best tradeoffs. This claim is not equally valid for general equilibrium analysis and intermarket second-best claims. If an affirmative defense for second-best concerns is successful in an intramarket setting, however, there is nothing to prevent broadening consideration to certain intermarket tradeoffs at a future date.

190. A theoretical concern still remains. This is the looming specter of the general theory of second best. Is it defensible to rely upon allocative efficiency as a basis for decision-making after acknowledging the significance of intramarket second-best tradeoffs? One of the primary lessons from Lipsey and Lancaster's seminal work is the caution against "piecemeal welfare economics." Lipsey & Lancaster, *supra* note 7, at 17. Given the existence of multiple market failures, correcting isolated market failures is not necessarily welfare-enhancing. Partial equilibrium analysis can start to unwind once the existence of market failures elsewhere in the economy is acknowledged. One could easily paraphrase Lipsey and Lancaster to say that in the presence of multiple market failures, there is no guarantee that maximizing total welfare on a market-by-market basis will necessarily lead to an increase in social welfare. My answer at this point will not be satisfying for those who seek to leverage Lipsey and Lancaster's insight into a full-scale reworking of contemporary antitrust analysis. *See supra* notes 10-11 and accompanying text (discussing Markovits's critique of allocative efficiency as a basis for antitrust analysis). In the end, my continued acceptance of allocative efficiency and the partial equilibrium framework as a basis for policymaking is premised on the belief that a properly constructed partial equilibrium analysis, and a decision rule that seeks to maximize total welfare, will capture most of the tradeoffs of first-order significance and will lead to more defensible results than the current antitrust rule that seeks either to minimize market power or to maximize "competition." This fact, in conjunction with the incorporation of dynamic efficiency concerns into the intramarket second-best defense and the express acknowledgment of an institutional partnership with the legislature to address intermarket tradeoffs, makes me comfortable with the limited second-best analysis advocated here.

Other defenses could also be mounted. Melvin Reder outlines the Chicago School's vision of general equilibrium analysis as a functioning research agenda. *See* Melvin W. Reder, *Chicago Economics: Permanence and Change*, 20 J. ECON. LIT. 1, 11-13 (1982). Core assumptions of research agendas are seldom questioned, at least by those inside the paradigm.

Finally, the antitrust enforcement agencies could play an important role in developing an intramarket second-best defense. Enforcers could acknowledge the relevance of second-best concerns, even if courts do not adopt the defense. Enforcement agency guidelines have been influential in shaping antitrust law in the past. The arguments developed in this paper suggest that the Department of Justice and the Federal Trade Commission should seriously consider incorporating second-best concerns in subsequent guidelines and should more expressly acknowledge the role that total welfare analysis plays in antitrust law. Moreover, just as the enforcement agencies have been on the vanguard of developing and implementing the tools required to assess the unilateral effects of merger, the agencies could use those same techniques to refine their assessments of the welfare implications of restraints more generally. This would further develop the economic techniques and sophistication required for implementing a second-best defense. In addition, the enforcement agencies would be wise to consider the implications of second-best theory in deciding how best — that is most in the public interest — to allocate their enforcement resources, and in deciding which cases to prosecute. Even a strong belief that antitrust law has a particular statutory vision of competition should not prohibit the enforcement agencies from considering total welfare effects in allocating their limited enforcement resources.

C. *Total Welfare: Moving Antitrust Law Beyond Competition*

A workable antitrust policy must be capable of functioning in the hands of enforcement agency officials deciding which antitrust problems to prosecute, of judges and jurors resolving antitrust disputes brought before them, and of lawyers counseling clients regarding what conduct is and is not likely to violate the antitrust laws. A total welfare standard provides a more workable antitrust orientation than

---

Within the Chicago School view, examples of market power and market failure are viewed as limited in scope. In the long run, they do not undermine the assertion that "prices of factors are good approximations to the opportunity cost of using them." *Id.* at 15. "Market failure or more generally failure of individual decision makers to achieve a Pareto-optimum, is treated like monopoly, an unusual situation, to be analyzed ad hoc but not requiring a shift of emphasis away from the basic competitive model." *Id.* at 16. From this perspective, correcting intramarket second-best tradeoffs could be an appropriate part of the research agenda, while addressing intermarket second-best tradeoffs would not. William Baumol mounts a more pragmatic defense. Baumol attempts to avoid the policy paralysis implicit in second-best theory by arguing that (1) many of the interdependencies may be sufficiently limited to permit effective partitioning of the market, (2) many policy problems are sufficiently pressing that some action must be taken even if there might be unforeseen and negative secondary consequences, (3) policies leading to welfare improvements may be justified even if they do not produce optimality, and (4) the economic theory of second best must contend with the political reality that comprehensive reforms are seldom politically available, forcing the need for a series of partial solutions. *See* Baumol, *supra* note 188, at 143-45. Still, Baumol cautions that policy advisers who casually ignore problems of second best do so "at the public's peril." *Id.* at 145.

does the existing competition-based approach. In making this assertion, I envision a total welfare standard as an evolutionary extension of the present competition-based framework. Most of the legal presumptions about anticompetitive conduct would be retained. The important difference would emerge in cases where economic facts call these presumptions into question.

Competition-based antitrust policies embrace structural understandings of what competition means and proceed to reject, as "anticompetitive," conduct that is inconsistent with that understanding. The problem with this approach is that competition is ill defined. Most understandings of competition (historical or contemporary) do not provide an apparatus sufficient to support a comprehensive antitrust policy. The classical conception of competition as rivalry, for example, falls short of the mark. It does not tell us how much rivalry is appropriate or what type of rivalry is desirable. In the name of rivalry should we allow competitors to torch each other's business? Moreover, the heart of any workable antitrust policy is the ability to strike an appropriate balance between cooperation and rivalry. An inherent limitation of competition-based orientations is that they address only one side of this equation.

Some courts have attempted to fill this void with a more developed neoclassical understanding of competition. An industry is perfectly competitive if there exist large numbers of buyers and sellers acting independently. "Large numbers" can be operationally defined by the inability of any single buyer or seller to affect market price (each producer faces a perfectly elastic firm-specific demand curve). One pitfall of this approach is the tendency to lose the larger picture and focus too closely on the effect the alleged anticompetitive conduct has on individual competitors. This pitfall can be avoided by repeating the admonition that antitrust laws are designed to protect competition and not competitors, and by recalling that the appropriate focus is not on the fate of individual firms, but rather on whether those firms remaining will have any ability unilaterally to influence price or output.

This insight reflects an important refinement of the structural understanding of competition. The presence of large numbers of buyers and sellers is important only as *evidence* from which it can be inferred that there is no market power. Competition can be understood as more functional than structural. It can be defined as the absence of producer discretion over price (reminiscent of the statement that producers in competitive markets are price takers and not price makers). If the focus is on the absence of individual firm discretion, then it makes sense to investigate not only the constraints imposed by existing competitors but also the constraints posed by potential competitors. From this perspective, it is possible to envision a market that is "competitive" yet has only a handful of functioning firms. The lesson is one that is already well recognized in antitrust law: large numbers

need not be the hallmark of competition, so long as there are no substantial barriers to entry.[191]

This is a richer and more textured understanding of how markets work, but the fact remains that even this more functional definition of competition breaks down as a guide to policy in important cases. These are cases in which cooperation is more beneficial than rivalry, and where economies of scale or scope are substantial. How should a competition-based antitrust policy address Williamson's merger that creates real market power and real welfare gains? There may be a small number of post-merger competitors, and the very nature of the technology generating the efficiencies may be the source of substantial barriers to entry. Alternatively, how should a competition-based policy address cooperative R&D joint ventures, or joint industry standard-setting in the face of imperfect information? When courts encounter these problems, they frequently slip into forms of analysis that are more consistent with a total welfare standard, even if they attempt to retain the rhetoric of competition.

Under contemporary doctrine, restraints of trade can be justified if the restraints are "procompetitive," but what does it mean to be "procompetitive"? Which justifications will this standard permit and which will it disallow? The answers that the Court often gives mark a departure from structural understandings of competition. Conduct is procompetitive if it increases output, reduces price, or enables the parties to provide a product or service that would not otherwise be available.[192] It is easier to reconcile these criteria with a total welfare standard than with a competition-based standard. Courts will permit conduct and agreements that might otherwise be viewed as restraints of trade, if the parties can establish that the conduct will result in a net increase in total welfare. Within this framework, evidence of an increase in output, a reduction in price, or the introduction of a new product most often indicates an increase in social welfare, regardless of the impact of the conduct on competition. By the same token, these factors do not constitute an exhaustive list of ways to demonstrate an increase in total welfare.

I do not advocate abandoning either the infrastructure of competitive analysis or the series of legal presumptions that have been built on its foundation. In a vast number of cases, a structural understanding of competition serves as an effective proxy for separating welfare-

191. *See, e.g.,* United States v. Waste Management, Inc., 743 F.2d 976 (2d Cir. 1984) (proof of ease of entry into a market can rebut a prima facie case of illegality based upon high market shares under section 7); *see also* cases cited *supra* note 171 (discussing barriers to entry).

192. *See, e.g.,* NCAA v. Board of Regents, 468 U.S. 85, 102-04 (1984) (defining "procompetitive effects" in terms of the ability to widen consumer choices, introduce a new product, or increase output); Broadcast Music, Inc. v. Columbia Broad. Sys., Inc., 441 U.S. 1, 18-23 (1979) (examining the "procompetitive" effects of BMI's blanket licenses).

enhancing (efficient) from welfare-destroying (inefficient) behavior. Furthermore, I generally favor heuristic devices and appropriate presumptions drawn from competitive theory. The presumptions of legality and illegality based upon market definition and concentration ratios in section 7 cases are generally defensible. Similarly, the presumptions of illegality for most of the categories of conduct considered per se illegal are also defensible.

I would build upon this structure by explicitly acknowledging that structural concepts of competition are used only as a proxy to determine the ultimate issue of whether conduct is efficient (welfare-enhancing). Hence, I would make these presumptions rebuttable — even in what now are per se cases. A party could rebut a presumption of illegality if it could demonstrate (by clear and convincing evidence) that the underlying conduct produced an increase in total welfare. In varying degrees, district and appellate courts are already doing this when they consider an efficiency defense in section 7 merger cases.[193] I believe that the Supreme Court is also implicitly doing this now in Sherman Act section 1 cases when it decides whether or not to invoke the per se rule or consider conduct under the rule of reason.[194] While preserving the rhetoric of competition, the Court is developing a substantive body of antitrust principles more consistent with a total welfare standard. A number of the examples discussed in Section I.B.3 illustrate how courts are addressing market failure concerns under the rule of reason. Once these trends are acknowledged, it is a short step to applying the same logic and analysis to intramarket second-best tradeoffs.

Taking this step beyond competition to express welfare analysis is important. A welfare orientation dovetails better with the reasoning underlying most private business transactions and lays the foundation for more effective client counseling. Often, the most effective question in antitrust counseling is, "Why do you want to do this deal?" or "What benefits do you expect and who will receive them?" While in practice, I felt comfortable telling clients that if they were confident that they could convince a group of lay outside observers that their conduct was "efficient" in the sense that it would produce net social

193. *See* cases cited *supra* note 172 (efficiency defense in section 7 cases).

194. *Compare, e.g.*, United States v. Topco Assocs., 405 U.S. 596, 608 (1972) (declaring per se illegal the territorial divisions underlying the efforts of small independent grocers to market their own private label brand), *with NCAA*, 468 U.S. at 100-04 (holding that the NCAA's collectively bargained television contract, which limited the aggregate number of games broadcast and prohibited member schools from independently marketing their own games, must be examined under the rule of reason); *compare also* Arizona v. Maricopa County Med. Soc'y, 457 U.S. 332, 342-55 (1982) (holding efforts by physicians to collectively negotiate terms and rates with third-party payors to be a per se illegal price fixing conspiracy), *with Broadcast Music*, 441 U.S. at 8-10, 24 (holding that the joint efforts of composers to negotiate blanket licenses was price fixing only in a "literal" sense and would be evaluated under the rule of reason).

benefits, then it was unlikely that the conduct would violate the antitrust laws. Although this welfare analysis serves as a good predictor of likely antitrust outcome, it is not a good predictor of the likelihood of a government challenge or the ultimate cost or length of litigation. Instead, the degree to which the conduct conforms with traditional conceptions of "competition" is a better predictor of the likelihood of challenge, the stage at which the dispute would be resolved, as well as the total litigation costs. Reforming antitrust law under a total welfare banner would help resolve this tension and realign antitrust outcomes, the likelihood of challenge, and the cost of proceedings in a more rational and consistent manner.

CONCLUSION

Intramarket second-best claims present an interesting puzzle: the paradox of "anticompetitive" conduct that may be welfare-enhancing. The problem highlights tensions between economic and noneconomic approaches to antitrust law, as well as tensions between different economic approaches, by driving a wedge between "competition" and "efficiency" as guiding principles. There are many parallels between second-best concerns and Oliver Williamson's productive efficiency defense. Thirty years ago, Williamson outlined a four-step process for the evolutionary adoption of an efficiency defense for horizontal mergers.[195] The first step was acknowledging efficiencies as a matter of principle, even if disallowed as a practical matter. The second step was the introduction of efficiency-related evidence for explanatory completeness, to understand the full context in which the merger takes place. The third step was litigating efficiencies as an actual defense in limited types of mergers. The final step was the introduction of an efficiencies defense more generally, as dictated by accumulated theory and practical experience. In this analysis, Williamson applied many of his own theories of institutional economics to the evolution and development of antitrust doctrine.[196] Three decades later, this analysis has proven to be prophetic. Accomplishing the first and perhaps second steps would be a substantial victory for intramarket second-best tradeoffs, presenting an opportunity for antitrust law to move beyond structural concepts of competition and to take the next logical step in its evolutionary development.

195. *See* Williamson, *supra* note 2, at 113-14.

196. *See id.*; *see also* Oliver E. Williamson, *Delimiting Antitrust*, 76 GEO. L.J. 271, 272-81 (1987) (outlining an evolutionary approach to the incorporation of economic theory into antitrust law that stresses the importance of legal process over legal rules in shaping antitrust doctrine).