On the Inefficiencies of Efficiency as the Single-minded Goal of Antitrust

Albert Foer*

Abstract
This article provides a skeptical look at the role of “efficiency” in antitrust analysis. It begins by discussing how various types of efficiency may be in conflict, such that the concept is actually less scientific or useful than often thought. The article asks, What economic considerations are excluded from today’s antitrust analysis? This leads to a discussion of political, social, and non-efficiency economic values; externalities; and three types of inefficiencies that may be caused by the single-minded hunt for efficiencies: X-inefficiency; diseconomies of scale, scope, and coordination; and the “too-big-to-fail” problem. Finally, the article describes some challenges facing reform proposals, including problems of prediction and quantification. The conclusion argues that the efficiency goal is fraught with problems and that a broader concept needs to take into account inefficiencies as well as efficiencies.

Keywords
antitrust, efficiency, X-inefficiency, diseconomies, consumer welfare

Many political words are similarly abused. The word fascism has now no meaning except in so far as it signifies “something not desirable.” The words democracy, socialism, freedom, patriotic, realistic, justice, have each of them several different meanings which cannot be reconciled with one another... Words of this kind are often used in a consciously dishonest way.

George Orwell†


*Founder and Senior Fellow, American Antitrust Institute, Washington, DC, USA

†Corresponding Author:
Albert Foer, Founder and Senior Fellow, American Antitrust Institute, Washington, DC 20008, USA; 2919 Ellicott Street, NW, Washington, DC 20008-1022, USA.
Email: bfoer@antitrustinstitute.org
I. Introduction

For the past quarter century, a full generation, one word has dominated the antitrust community as the castle dominates Edinburgh or the Parthenon, Athens. The word is efficiency. In his 2001 edition of Antitrust Law, Richard Posner, once my antitrust professor at the University of Chicago and now our most celebrated federal Court of Appeals judge, wrote:

Almost everyone professionally involved in antitrust today—whether as litigator, prosecutor, judge, academic, or informed observer—not only agrees that the only goal of the antitrust laws should be to promote economic welfare, but also agrees on the essential tenets of economic theory that should be used to determine the consistency of specific business practices with that goal. Agrees, that is, that economic welfare should be understood in terms of the economist’s concept of efficiency. ²

This goes further than William Baxter went in 1983, when he recognized that conflicting views of the goals of antitrust still had to be taken into account, even if subordinated to efficiency:

[W]here there is a conflict, social and political goals should yield to economic considerations primarily for two reasons: first, the statutes themselves focus on efficiency; and second, nonefficiency goals are too intractable to be used as enforcement standards.³

Notice that Baxter initially distinguished social and political goals from economic considerations, but quickly jumped to implying that non-efficiency goals are necessarily non-economic in nature. This is consistent with the position of Frank Easterbrook, my classmate in Professor Posner’s antitrust class at Chicago, who urged the efficiency über alles doctrine despite its departure from the historic origins of antitrust. Easterbrook, now serving on the same appellate court as Posner, wrote in 1981:

I agree with Robert Bork that, whatever one makes of this history, the antitrust laws should be treated as if they served no goal other than economic efficiency.⁴

That the twenty-first-century Posnerian consensus reflects a growing self-confidence on the part of Chicago School adherents is emphasized by the absence of his triumphal 2001 observation in the first edition of Posner’s Antitrust Law in 1976, whose purpose was then merely to demonstrate the importance of economic analysis to antitrust.⁵ The difference between promulgation of an idea and assertion a quarter century later that the idea had become virtually unchallenged is rather breathtaking. I believe it claims too much.⁶

5. In his first edition, Posner wrote, “The work of the economists provides at least a starting point for analysis. Since, unfortunately, they are not unanimous on the essential points of the theory of monopoly, a necessary first step is to thread one’s way through the doctrinal controversies that have surrounded and continue to afflict the development of the theory.” Richard A. Posner, Antitrust Law: An Economic Perspective 3–4 (1st ed., 1976).
6. The Organization for Economic Co-operation and Development in 2012 conducted a policy roundtable, The Role of Efficiency Claims in Antitrust Proceedings (2012), http://www.oecd.org/daf/competition/efficiencyclaims2012.pdf, the report for which offers a great deal of background on the subject of efficiency, including a review of how attitudes toward efficiencies have changed over time in the United States, the European Union, and in other jurisdictions.
A recent proposal by Federal Trade Commissioner Joshua Wright that the Federal Trade Commission (FTC) enact guidelines for the interpretation of Section 5 of the FTC Act has created controversy by raising the efficiency mantra to a new peak.\(^7\) His central test of illegality is whether a practice "generates harm to competition as understood by the traditional antitrust laws and generates no cognizable efficiencies."\(^8\) In effect, his proposal would say that a stand-alone Section 5 case (i.e., an FTC complaint not also relying on the Sherman Act or the Clayton Act, but only on the FTC Act's prohibition of "unfair methods of competition") could not reach acts or practices that create any quantum, however small, of cognizable efficiency. In effect, this would translate the simple Section 5 proscription of unfairness in competition into something akin to "methods of competition that fail to create even an iota of efficiency"—thereby completing the Chicago School's capture of the competition flag.

Note that this proposal is not about weighing harm to competition against efficiency gains; it is not even about weighing the gain to efficiency. It is more like the old Groucho Marx television quiz show, You Bet Your Life: if the contestant accidentally utters the night's magic word, for example "efficiency," a duck will fly down and deposit a munificent sum of money. Without necessarily attributing political dishonesty to anyone, I believe George Orwell, were he writing about antitrust today, would add the word "efficiency" to his vocabulary of abused political words.

However this magical word "efficiency" may be defined (and that topic will be taken up shortly), there is general consensus, in the words of Professor Eleanor Fox, that "We want an efficient, growing economy. We want markets with efficient, creative, inventive firms."\(^9\) But do we know how to get there from here? Where does antitrust fit into the larger economic picture? How can it help us achieve our shared goals? Since the 1980s, says Professor Fox, it has been assumed that the proper path is to let "the market" work and to avoid governmental intervention unless it can be demonstrated that the transaction or conduct is inefficient in a specific market-power-creating way. This has led, not accidentally, to high levels of concentration in many markets.

A new conversation is essential to consider whether there are better ways to reach the common objectives. To paraphrase Professor Fox,\(^10\) perhaps a better focus would be on preserving the autonomy of rivals and potential rivals, preserving abilities to adjust and adapt, and keeping the field clearer for competition on the merits by those without power.\(^11\) How can these objectives be reconciled with a high regard for efficiency?

I believe we have allowed consolidation in the name of efficiency to go too far, endangering the dynamic efficiency that sometimes can in some sense be inefficient in the near term but offers a better chance for economic success in the longer term. We have learned a lot since the days of Alfred Marshall about dynamics and the importance of diversity, both in nature and economics, but it is not clear

---

8. Id.
10. Id.
11. In recent case law, the European Court of Justice (CJ) has made the same point. The CJ decisively rejected an efficiency-only approach and reiterated that the protective purpose of EU antitrust law is much wider, stressing the continuing relevance of protecting the competitive process as such. Joined Cases C-501/06 P, C-513/06 P, C-515/06 P, and C-519/06 P, GlaxoSmithKline Services Unlimited v. Comm'n (2009), E.C.R. I-9291, para. 63 ("[I]t must be borne in mind that the Court has held that, like other competition rules laid down in the Treaty, Article 81 EC aims to protect not only the interests of competitors or of consumers, but also the structure of the market and, in so doing, competition as such. Consequently, for a finding that an agreement has an anticompetitive object, it is not necessary that final consumers be deprived of the advantages of effective competition in terms of supply or price"). See also Ben Van Rompuy, Economic Efficiency: The Sole Concern of Modern Antitrust Policy? Non-efficiency Considerations under Article 101 TFEU (2012), at 198–99.
that we know how to engineer efficiency that is both lasting and consistent with our political and social values. In a democratic society, however, there is a lot to be said not only for separation of powers, checks and balances, and similar decentralizing, and hence inefficient political strategies embedded in our Constitution, but also for the leavening of economic diversity and serendipitous rivalry through competitive markets.

In contributing to the conversation that I am attempting to invoke, the questions I want to pose in 2015 amount to a skeptical re-examination of whether the self-confidence of the Chicago School in regard to efficiency is justified. What do we mean by “efficiency”? Is it really as scientific and objective as it sounds or does it, in Orwell’s words, have “several different meanings which cannot be reconciled with one another”? What do we mean by “economic analysis”? Can there be credible economic analysis that is not strictly limited to efficiency? Putting theory aside, what role does efficiency actually play in the investigation, litigation, and remedial phases of antitrust? Without doubting that economic analysis should be one central focus of antitrust enforcement and without re-arguing either the legislative history of antitrust laws or the interpretation of these laws at various times in our history, I intend to focus primarily on the economic (rather than social or political) factors that seem to be excluded by what has become the mantra of efficiency. I will examine whether a single-minded drive for “efficiency” masks losses to other types of economic values, resulting in a skewing of economic analysis and policy outcomes. Finally, I will comment on alternative ways for thinking about inefficiency within the antitrust context.

II. What do We mean by “Efficiency”?

A. The Varieties of Efficiency: Allocative, Productive, and Dynamic Efficiencies

If you look up the definitions of “efficiency” and “efficient” in my old edition of The Oxford Universal Dictionary, you won’t find anything particularly relevant to what today’s antitrust community discusses. Efficiency was no more than “the fact of being an efficient cause (now only in philosophical use),” and “efficient” meant no more than “productive of effects; operative.” Webster’s New Collegiate Dictionary comes a little closer in its 1976 definition, viz.: “effective operation as measured by a comparison of production with cost.”

Turning from old dictionaries to a popular current text on industrial organization economics, we come on Dennis W. Carlton and Jeffrey M. Perloff, who advise:

The standard assumption in most economic models is that the primary objective of a manager of a firm is to maximize the firm’s profits. The manager must sell the optimal amount of output, and the firm engages in efficient production: No more output could be produced with existing technology, given the quantity of inputs used.

14. WEBSTER’S NEW COLLEGIATE DICTIONARY 362 (1976 ed.).
15. DENNIS W. CARLTON & JEFFREY M. PEROFF, MODERN INDUSTRIAL ORGANIZATION 12 (4th ed. 2005). We could, but will not, segue here to a discussion of the foundations of welfare economics, Pareto optimality, the Kaldor-Hicks variation, theory of the second best, and so on, which attempt to establish ways to compare claims of efficiency. See Albert A. Foer, The Goals of Antitrust: Thoughts on Consumer Welfare in the US, in HANDBOOK OF RESEARCH IN TRANS-ATLANTIC ANTITRUST 566, 579–81 (Philip Marsden ed., 2006). This subject is not taken up in Carlton and Perloff, id., and can remain outside of our concerns in this article.
Carlton and Perloff write, “Most economists believe that the antitrust laws should have the very simple goal of promoting efficiency.” Even if this proposition is accepted, they go on to admit, “[E]conomists often have difficulty determining which practices result in inefficient behavior.” Presumably a similar difficulty would apply to determining efficient behavior.

Part of the difficulty lies in the fact that efficiency itself is far from a simple concept and is susceptible to multiple meanings. In the Carlton/Perloff quote, the reference is to efficiency in production. Productive efficiency refers to a firm’s production and distribution at the lowest feasible cost. But there are at least two other categories of efficiency that are broadly recognized by economists. Allocative efficiency refers to the allocation of resources to their most valued uses. And dynamic efficiency refers to the efficiency benefits achieved through research, development, and innovation.

In this article, I will talk not only about efficiency, but non-efficiency economic considerations and non-economic considerations. These terms are sometimes used in the literature interchangeably with “public interest or public policy” considerations and “non-competition considerations.” I make a distinction between non-efficiency considerations and non-economic considerations on the basis that non-economic considerations can include values such as ethical arguments against market behavior or political theories that have little if anything to do with economics, whereas non-efficiency considerations can include arguments that are economic in nature but that may have relatively little to do directly with efficiency. In any event, efficiency and non-efficiency considerations are not necessarily inconsistent; as Van Rompuy observes, they can be mutually enforcing. (“Considerations” in this context, he notes, are factors considered in a judgment or decision and are different from objectives of antitrust policy.)

B. Conflict in Priorities Among Efficiency Types

These three categories of efficiency may or may not lead economists to the same analytical outcome when they look at the facts (which may actually be forecasts, i.e. predictions) in a particular antitrust case. Productive efficiency seems to be a necessary condition for allocative efficiency, but allocative efficiency does not necessarily result from productive efficiency, as the example of monopoly, with its dead weight loss, may demonstrate. And it is easy to imagine a merger or other types of conduct within the marketplace that creates near-term static efficiencies, but which also diminishes the likelihood of the future benefits of innovation that would represent dynamic efficiency. Further, it is also possible to imagine giving so much credibility to predictions of large dynamic efficiencies to be realized at some relatively distant point in the future, that dramatic gains of efficiency in the nearer term are foreclosed. Because of the potential for conflict among the three types of efficiency, antitrust decision making sometimes requires a ranking and weighting of the three efficiencies, although this has only rarely been done explicitly.

The prevalent view seems in practice to prioritize the three categories in this fashion: the most important goal is allocative efficiency (focusing on profitability and price/cost margins), followed by productive efficiency (focusing on reducing costs), followed by dynamic efficiency (focusing on

18. I am using definitions in Eleanor Fox, The Efficiency Paradox, in How the Chicago School Overshot the Mark 77, 78 (Robert Pitofsky ed., 2008). A classic text in the field is Burton H. Klein, Dynamic Economics (1977). A fourth type of efficiency, consumption efficiency, is cited in Andrew J. Gavil, William Kovacic & Jonathan Baker, Antitrust Law in Perspective: Cases, Concepts and Problems in Competition Policy 28–29 (2002), and defined there as assuring that buyers who value the goods the most get them. This strikes me as already implied in the concept of allocative efficiency.
19. Terminology is discussed in Van Rompuy, supra note 11.
20. Id.
21. Id.
innovation effects). This ordering, reflecting a kind of prudence that values a single bird in the hand more than two or more that may or may not fly out of the bush, has been challenged by the eminent business consultant (and economist) Michael Porter, among many others. Recognizing that technological breakthroughs add more to economic growth than static improvements in efficiency, Porter would rank innovation as the highest goal, followed by what he calls value improvement (i.e., static production efficiency), leaving profitability/price-cost margins (allocative efficiency) at the gate. I would speculate that Porter’s priorities reflect his early awareness of the global and technological trends that inform movement toward rapid, microchip-led networks, systems, and platforms that increase the importance of innovation as a form of competitive advantage.

The relative value placed on the three efficiencies may be affected by the state of the economy at a particular time. For instance, Michael Perelman has generalized that when the economy is contracting, businesses tend to focus on operating more efficiently by cutting costs, while in a growing economy, the focus tends to be on investment (adding capacity), which implies that expectations for the future take precedence over dealing with the present. In other words, static efficiency may be more closely associated with a stagnant or declining economy while dynamic efficiency is more associated with a growing economy. Moreover, one’s ranking of the three types of efficiency may vary with the position or even the temperament of the one doing the ranking.

The potential for conflict among the three forms of efficiency is seen in the following quote from Jerry Ellig’s introduction to a compilation of essays titled Dynamic Competition and Public Policy:

The critical antitrust issue is not just whether a particular exclusionary practice produces some identifiable consumer benefit in the present, but also how that practice will affect the path of innovation in the future.

Stated differently, a focus on costs and prices—the neoclassical foundations for static efficiencies—may give a satisfactory near-term result, but this must be compared with what may be an ultimately more important long-term result that could be achieved by a different analysis.

The point is that these three kinds of efficiencies do not necessarily lead in the same antitrust direction, and different experts may assign different priorities when there is conflict. Designating the goal of antitrust as “efficiency” therefore may signify a false consensus in much the same way that we all may affirm our support for the triple political goals of the French Revolution, “Freedom, Fraternity, Equality,”—until we unpack the bundle and discover that you can’t necessarily have all three at the same time in equal measure. But is the word “efficiency” no more than the opposite of Orwell’s “fascism,” which he says simply signifies “something not desirable”?


26. See the recent opinion in Saint Alphonsus Medical Center–Nampa, Inc. v. St. Luke’s Health System, No. 1:13-CV-00116-BLW (D. Idaho Jan. 24, 2014), http://www.ftc.gov/system/files/documents/cases/140124stlukesfindings.pdf. There, the court accepts the Merger Guidelines as the basis for legal conclusions and lays out a variety of efficiencies that everyone on both sides recognized, but in the end, finds them not to be “merger specific” (i.e., the efficiencies could have been gained without the merger), and therefore rejects the merger.
III. The Breadth of Economics

A. Who is an Economist?

Names and labels are important to the story of efficiency in antitrust. Since the 1980s, antitrust has been dominated by a branch of academia associated with what is usually called the Chicago School. When I speak of the Chicago School, I am not thinking of the long and glorious tradition characterized by Johan Van Overveldt as reflecting “a strong work ethic, an unshakable belief in economics as a true science, academic excellence as the sole criterion for advancement, an intense debating culture focused on sharpening the critical mind, and the University of Chicago’s two-dimensional isolation.”27 Rather, I am thinking of the antitrust theories that grew in popularity during the 1970s and ’80s and are associated with Aaron Director, George Stigler, Richard Posner, Robert Bork, Ronald Coase, Frank Easterbrook, and others sharing the basic neoclassical assumptions: viz, that (1) price theory is the supreme economic model, (2) markets rarely fail, (3) economic man is a rational decision maker whose goal is to maximize profit, and (4) government can rarely do anything that actually serves the public interest when it regulates economic activities.

When the Chicago School itself refers to economics, it does not speak of the broad range of professionals who have been trained in university economics and who think of themselves as economists. Rather, it seems to arrogate the entire economic realm unto itself. This is an act of legerdemain, similar to Robert Bork’s capture of the term “consumer welfare” for the Chicago School’s flavor of antitrust, portraying a concept of total efficiency for the economy (including producer surplus) as if it is of particularized interest to consumers.

While the Chicago School’s views have unquestionably been influential, even with many students of antitrust who do not affiliate with Chicago, many neoclassical economists have criticized the Chicago School for some of its positions. Outstanding economists such as F.M. Scherer, Walter Adams, and William Comanor leaned strongly against the Chicago tide, in the tradition of the so-called structure-conduct-performance school of economics often associated with the earlier work of Edward Mason, Joe Bain, and others trained at Harvard University.28 During the past twenty years, many professionals with economics training, often with a strongly empirical bent, criticized the Chicago School approach as too simplistic. A “post-Chicago” economics gained acceptance “as an intellectually rigorous alternative approach to antitrust.”29 In particular, these critics both in economics departments and in business and law schools tend to argue that the strategic management of business deserves a larger role in explaining how the economy in fact operates. Many economically sophisticated MBAs profess in business schools where their empirically informed understanding of business behavior augments or sometimes conflicts with more theory-based economic experts. I will write more about this in the next section.

In fact, beyond the neoclassical realm, other economists venture. Behavioral economists, taking their cues from psychology (many of whom are, ironically, also associated with the University of Chicago) undermine the Chicago model of the rational economic man. These economists are given respect but told that their insights do not matter for antitrust because, on the whole, the rational man model works well enough. (The Ptolemaic system also worked well enough for certain astronomical

calculations.) Meanwhile, the house of economics continues to grow and should be encouraged, where intellectually appropriate, to influence thinking about antitrust. Institutional economists, not surprisingly, stress the role of institutions (individuals, firms, states, social norms), and evolutionary economists stress the role of long-term evolutionary factors in shaping economic behavior. As the Chicago School promoted its demand that economic effects be the sole test of antitrust, economists who emphasize structural components in industrial organization are denounced as antibusiness formalists. Populist economists who give a higher value to the political goals of regulation than to neoclassical theorizing are dismissed as atavistic throwbacks to a bygone era. The extent to which these non-Chicago economists will influence antitrust remains to be seen. It simply is not true, however, that one must swear allegiance to the Chicago School in order to be recognized as a competent economist.

I often talk about resurrecting the term “political economy” as a modern label for the mix of disciplines that can counter Chicago’s narrowness. A sidebar on this, as described by Michael Perelman, is interesting. At an earlier stage in American history, what we now call “economics” was called “political economy” and was synonymous with laissez faire. A part of the political economy professoriate that represented the profession’s corporatist wing did not think laissez faire theory was applicable to a new industrial age in which large amounts of fixed capital could not be recovered under a regime of competition. In the time of many railroad bankruptcies and subsequent consolidations, where pricing at the margin left no room to recover large fixed capital, they were convinced that competition was inefficient, standing in the way of large-scale investment in firms of large scale. This group, in 1885, spearheaded creation of the American Economics Association for the purpose of separating corporatist “economics” from laissez faire “political economy.” Ironically, whereas the earlier laissez faire theories—resting on an economic base of small enterprises and fragmented competition—recognized a relationship between politics and economy, the new American Economics Association created by corporatists of a big-business mind-set was rather quickly taken over by the laissez faire advocates, who today purport to be the acolytes of the science of depoliticized economics. Be that as it may, today it is those who criticize the Chicago School’s laissez faire theories who want to bring back the label of “political economy” to distinguish a broader view of economics that includes non-efficiency values.

Political scientist Marc Allen Eisner describes the revolution in antitrust that occurred in the 1980s in a book titled Antitrust and the Triumph of Economics. He begins by describing “the broad bipartisan consensus that protected antitrust for much of its history” and how it had disintegrated by the 1980s. He writes:

The Reagan administration’s Chicago school interpretation of antitrust deemphasized and in many cases rejected the political and social goals of policy. The only objective that could be correctly derived from

30. “In academia, the shift has come partly through the introduction of new tools, and models that reveal the shortcomings of unfettered capitalism. Game theory shows how competition can lead to waste, and models of asymmetric information also show how markets can fail. Decision theory, learning theory and behavioral economics have poked holes in the old assumptions of perfect rationality. Even in macroeconomics, all the focus is on incomplete and imperfectly functioning markets, as Karthik Athreya explains in his recent book, Big Ideas in Macroeconomics. . . . The move away from pure free-marketeerism has been helped by a flood of new data. Economics has become much more empirical, and that has made it much harder to wave away the possibility of market inefficiencies.” Noah Smith, Economics Stars Swing Left, Bloomberg View, Jan. 7, 2015, http://www.bloombergview.com/articles/2015-01-07/economics-stars-swing-left.
31. PERELMAN, supra note 24, at 101–03.
32. Id. at 103.
34. Id. at 2.
the history of antitrust was the promotion of business efficiency—a goal often best realized through systematic deference to the market.\textsuperscript{35}

Eisner describes how during the transition from the old antitrust to the new antitrust, influence within the antitrust enforcement agencies shifted from lawyers (with their natural emphasis on process and legislative intent) to economists. In fact, economists had long played a real, albeit secondary, role in the development of antitrust enforcement policy, generally teaching a paradigm that related market structure to firm conduct to market performance.\textsuperscript{36} Nonetheless, Eisner’s description of this transmogrification of antitrust’s occupational sociology is largely valid. Lawyers did not go away, of course; rather, they became compelled to pay more attention to the economists and their theories—indeed, to speak the language—as the analytical emphasis moved from emphasizing industrial structure to determining competitive effects. The legal and economic analyses consequently became far more integrated. But here is what is important: the matter was not simply that economists became more influential in the academy and in the antitrust bureaucracy; it is that the Chicago School became the politically dominant wing of the economics profession in the field of antitrust. This was not merely the triumph of economics, so much as the triumph of a particularly conservative brand of economics.

B. Business Schools versus Economics Departments

In 2002, the American Antitrust Institute investigated the relationship between what the Chicago School was teaching economists about competition and what business schools were teaching future corporate executives about competition.\textsuperscript{37} The results were somewhat surprising. Whereas Chicago School economists tended to focus on theoretical models based on assumptions about how people act, the business schools tended to focus on empirical understanding of what in fact works in the marketplace—not the marketplace of ideas but the marketplace of commerce. We found that competition tends to be covered in academic MBA classes by three separate faculty groups within business schools: the marketing faculty, the strategic management faculty, and the economics faculty.\textsuperscript{38}

While the economics faculty in a business school tends to teach the same brand of economics as would be taught in a graduate economics department, with perhaps more of an applications orientation,\textsuperscript{39} the marketing and strategic management faculties tend to deviate in at least three substantial ways from the Chicago School models.

\textsuperscript{35} Id. at 3.


\textsuperscript{37} The study was presented at an AAI conference in 2002, which generated papers included in a symposium volume, \textit{The Dialogue Between Students of Business and Students of Antitrust}, 47 N.Y.L. SCH. L. REV. 1 (2003).


First, whereas the Chicago School assumes the model of a rational economic decision maker, the marketing faculty takes into more prominent account the systematic deviations from rationality that have come to be associated with behavioral psychology and behavioral economics: all is not necessarily rational, and irrationalities can play a role in thinking about how to compete. Advertising, for example, is often targeted at consumer emotions rather than at cold calculation of costs and benefits. An increasing amount of attention is also being given to the extent that firms, like individuals, are subject to patterns of irrational behavior.40

Second, the Chicago School tends to assume that the market works with full information, whereas the business schools recognize the need to make decisions in the absence of full information, often in the fog of competition. Thus, what may appear to be efficient could be based on incomplete information.

And third, the Chicago School assumes that the goal of business decisions is to maximize profit for the firm, whereas the business schools tend to teach that it is difficult to know what strategy or price will maximize profit, and that a more realistic goal is to establish and maintain a competitive advantage as the pathway to long-term success that includes profits. In other words, market power, which is neither necessarily bad nor necessarily illegal, often surpasses profit maximization as a motivating factor.

These significant differences between what the Chicago School was teaching to economists and what the business schools were teaching to future executives suggest a disconnect that has implications for the role of efficiency in antitrust. Although MBAs are not necessarily to be accredited as economists, they are typically taught a good deal of economics, and their insights into (and influence on) actual business behavior would seem highly relevant to public policies based on “the real world.” In the decade since our study, the gap may have closed somewhat, especially with the increased authority that behavioral economics has gained, for example, through award of the Nobel Prize in Economics to Daniel Kahneman,41 the increased attention being given by economists to the role of strategy in business decision making,42 and the increased attention being paid to sustainability and the responsibilities of management to multiple stakeholders.

My points here are simple: (1) economics is a much broader field than Chicago School economics, and (2) the current elevation of efficiency in antitrust analysis should not be taken as universally accepted by the full community of experts that pay close attention to competition policy and antitrust.

IV. The Role of Efficiency in Antitrust Analysis

A. Federal Merger Guidelines

During my days as a Senior Executive of the FTC in the mid-1970s, people were beginning to discuss whether efficiencies should be taken into account during enforcement proceedings. Staff’s unofficial response tended to be something like: “Of course we take efficiency into account as part of our prosecutorial discretion, but it would require a very strong showing of very strong efficiencies to overcome our duty to prosecute anticompetitive activities.” Moreover, we opposed formalizing a role

---

42. CARLTON & PERLOFF, supra note 15, ch. 11 (on strategic behavior).
for efficiencies, anticipating that this would make trials too complicated and the outcome too unpredictable. “Just leave the subject to our prosecutorial discretion,” we argued.

Before the ‘80s, antitrust was largely about protecting open markets and the competitive process. A marker in the shift away from this focus was the way that efficiency gained a foothold in the Horizontal Merger Guidelines jointly issued by the Department of Justice (DOJ) and FTC. The first guidelines, introduced by the Reagan administration in 1982, stated that all mergers resulted in some modest degree of efficiencies, so that only in cases involving “unusual” (that is, atypically large) efficiencies would efficiency receive specific attention. But the 1984 revision said the agencies would consider merger-related cost savings in all cases, which of course encouraged all merging parties to submit claims of efficiencies.

This gave prominence to a new inquiry: “Will the outcome of a particular merger or conduct be inefficient by inducing the aggregate of all producers to reduce the total amount of goods they produce?” Posing the issue this way seemed to assume the almost inevitability of efficiency gains in horizontal mergers and placed the burden on enforcers of antitrust laws to demonstrate inefficiency. The 1997 amendments to the Merger Guidelines provided “the most complete statement concerning the role of efficiencies ever to come from the federal enforcement agencies.” (These became the template for the Collaboration Guidelines issued in 2000 to cover joint ventures.) The Merger Guidelines were again amended in 2010.

Section 10 of the Merger Guidelines deals with efficiency. It does not actually define efficiency, but offers examples. Only certain efficiencies are deemed cognizable. This key concept holds that to be counted, an efficiency must be both merger-specific (i.e., not achievable without this merger) and verifiable. The examples show that the Merger Guidelines focus primarily on the production of lower costs in the product development, production, or marketing of the merging or collaborating firms.

Despite the ever-increasing formal role of efficiency in merger analysis, the Merger Guidelines reflect the agencies’ abiding and generally skeptical attitude toward efficiency claims. For example,

The greater the potential adverse competitive effect of a merger, the greater must be the cognizable efficiencies, and the more they must be passed through to customers, for the Agencies to conclude that the merger will not have an anticompetitive effect in the relevant market.

While the Merger Guidelines have brought welcomed structure to the analysis of mergers, they have, inevitably, not been able to answer all questions. For example, in regard to efficiency projections, what role should be assigned to efficiencies claimed for an overall national system or for markets outside of the markets in which anticompetitive effects will occur? The somewhat infamous footnote 14 leaves

43. Eleanor Fox, supra note 18, at 79.
44. Andrew L. Gavel et al., supra note 18, at 890.
46. “It is incumbent upon the merging firms to substantiate efficiency claims so that the Agencies can verify by reasonable means the likelihood and magnitude of each asserted efficiency, how and when each would be achieved (and any costs of doing so), how each would enhance the merged firm’s ability and incentive to compete, and why each would be merger-specific.” HORIZONTAL MERGER GUIDELINES, supra note 45, at § 10.
47. Id.
this to prosecutorial discretion, which is to say, without transparency or predictability.\textsuperscript{48} A recent instance is the remedy agreed to by the DOJ and two merging airlines, American Airlines and US Airways, where anticompetitive effects alleged by the DOJ in many markets were arguably ignored in the settlement in favor of systemwide efficiencies predicted on the basis of a small number of divestiture of slots in a few airports to two low-cost carriers.\textsuperscript{49}

**B. Differences Between Mergers and Other Conduct**

Efficiency is built into the Merger and Collaboration Guidelines because both mergers and joint ventures can be defended in terms of their ability to create new efficiencies that are not available if the firms remain fully independent. The legal question under the federal guidelines is whether a proposed merger or collaboration should be permitted. That requires determination whether the consolidation will tend to create anticompetitive restraints on trade in the future and, if so, whether the society will benefit so much from the anticipated efficiencies created by the merger that whatever anticompetitive harm there may be should be considered outweighed. In short, the enforcers and ultimately the courts are being asked to make predictions about future efficiencies, even before they come to a weighing of what may be incommensurables.\textsuperscript{50}

Courts have generally followed the federal Horizontal Merger Guidelines in their analysis of contested mergers. Relatively few cases have focused on efficiency claims.\textsuperscript{51}

Although there may be elements of prediction in other types of conduct alleged to violate the antitrust laws, typically the alleged harm will already have occurred and any evaluation of efficiency will be founded on actual rather than predicted facts. Rather than having a formalized role for efficiency as a defense, non-merger cases will consider efficiency arguments, if at all, in the context of a Rule of Reason analysis, but then only in the context of whether the challenged action advanced or suppressed competition. The Supreme Court in *National Society of Professional Engineers* held:

\textsuperscript{48} Footnote 14 provides: “The Agencies normally assess competition in each relevant market affected by a merger independently and normally will challenge the merger if it is likely to be anticompetitive in any relevant market. In some cases, however, the Agencies in their prosecutorial discretion will consider efficiencies not strictly in the relevant market, but so inextricably linked with it that a partial divestiture or other remedy could not feasibly eliminate the anticompetitive effect in the relevant market without sacrificing the efficiencies in the other market(s). Inextricably linked efficiencies are most likely to make a difference when they are great and the likely anticompetitive effect in the relevant market(s) is small so the merger is likely to benefit customers overall.” *Id.* at 30 n.14. It is argued by an FTC Commissioner and his Attorney Advisor that out-of-markets efficiencies should be required to be considered in mergers. Jan M. Ribniecek & Joshua D. Wright, *Outside In or Inside Out? Counting Merger Efficiencies Inside and Out of the Relevant Market*, in 2 WILLIAM E. KOVACK, AN ANTITRUST TRIBUTE 443 (Nicolas Charbit & Elisa Ramundo eds., 2014).

\textsuperscript{49} See the critical comments on *United States v. US Airways Group, Inc.* from American Antitrust Institute and others to William H. Stallings, Chief, Transportation, Energy, and Agriculture Section, Antitrust Division, U.S. Department of Justice (Feb. 7, 2013), http://www.antitrustinstitute.org/content/aii-objects-airline-merger-settlement.


The assumption that competition is the best method of allocating resources in a free market recognizes that all elements of a bargain—quality, service, safety, and durability—and not just the immediate cost, are favorably affected by the free opportunity to select among alternative offers. Even assuming occasional exceptions to the presumed consequences of competition, the statutory policy precludes inquiry into the question whether competition is good or bad.  

The immediate effect of this decision was to say that constraints on when price competition could occur in the course of the sale of engineering services, set out in a professional organization’s code of ethics, would be judged on a per se basis as price fixing, and that arguments about serious negative consequences, such as bridges being more likely to collapse, could not be considered. I will return to this in a few pages.

C. Remedies

Beyond questions of prosecutorial discretion in the investigation phase of a possible antitrust case, and beyond the question of how much weight to give to efficiency in the adjudication of whether the law has been violated, consideration of efficiency may arise importantly after liability has been determined, as a remedy is fashioned. Here, again, the focus is necessarily on the future, and it might well be argued that this is the area in which the broadest concepts of efficiency and inefficiency might be applied because discretionary judgments play a very large role and the impact of a remedy may play out over many years, with consequences in many markets.

D. The Need for Balance

The legal/economic advocate’s mind is far less constrained when predicting the future than interpreting the past. In dealing with the past, certain possibilities will be foreclosed, because of what did or did not happen. Evidence must be explained. Claims about future efficiencies or inefficiencies, on the other hand, can be more creative and far-reaching, and less subject to proof or disproof than past facts. The question of burden of proof with respect to the future is, in a sense, irrelevant because nothing can be proved about the future. It is possible only to clarify assumptions and offer probabilities of various scenarios coming true at some point or points in the future. This holds both for claims of future efficiency and claims of future inefficiency. To the extent such claims are to be considered, the decision maker must take into account both probability that the claimed efficiency or inefficiency will actually occur and the magnitude of the effect if it does occur.

Administrability of the system and justiciability become important considerations in this situation. The more factors that are deemed cognizable, the longer it will take to resolve conflicts (giving advantage to the party that benefits from delay); the more expensive will be a trial (giving advantage to the deeper pockets); the more complicated the issues, the greater the discretion will be lodged in the decision makers (i.e., the prosecutors and judges); and the less transparent and understandable will be the outcome to the people who ultimately have to decide whether they support antitrust as a political mechanism (the public and its elected representatives). Perhaps most importantly, the party having to carry the burden of proof is more likely to lose.

It is therefore important to find the right balance in antitrust between the search for something called efficiency and the administrability and justiciability of an antitrust system. Given the difficulty, if not the logical impossibility, of proving the future, if the government had to prove the absence of future efficiencies, virtually any expert’s claim would justify approval of whatever it is that the expert’s client wants to do. The public interest in regulating anticompetitive conduct would be overrun. On the other

hand, the idea that all arm's-length business transactions are likely to involve some efficiencies, for which the antitrust laws give a "standard deduction," and that only remarkably large efficiencies will be explored and weighed as a possible defense, with the burden on the parties seeking to justify their claim, seems to establish a reasonable balance.

V. What is Excluded from Today's Antitrust Analysis?

In those situations where efficiencies are to be taken into account, it is obviously important to know what this means. We approach this, however, by asking what is excluded from the weighing.

A. Political, Social, and Non-efficiency Economic Values

The way that antitrust analysis has developed, a rather extensive set of blinders is used to keep the discussion narrowly focused on competitive effects. In some cases, the blinders take the form of evidentiary rules, so that, for example, the per se analysis of cartel price fixing skips over the question of whether there are either negative effects or positive efficiency benefits, going directly to the question of whether prices or output were restrained by agreement. In other cases, there are procedural rules that eliminate consideration of certain effects, such as the standing limitation that keeps a competitor from arguing that a company's increased efficiency makes the challenger less able to compete. In still other cases, negative effects on the economy or important values such as safety, reliability, or environmental health are precluded from consideration.

Because Section 7 of the Clayton Act bases illegality of a merger on the likelihood of anticompetitive effects in "any section of the country," some efficiencies that may support a merger are blocked by a tradition that precludes benefits from being considered if they are outside of the market in question. (Merging parties generally cannot argue that although the impact will be anticompetitive in Market A, the new company will improve competition in Market B. 53)

In general, political and social values such as the preservation of small businesses or the decentralization of political power through economic deconcentration are today not considered valid justifications for antitrust enforcement. The justification for exclusion has generally been that political/social values are too subjective and that their incorporation into antitrust analysis would inevitably lead to unpredictable rules that could change with each election, and decisions might also be subject to corruption. However, economics is a social science, not a natural science and definitely not a subdiscipline of mathematics; the line between what is political/social and what is economic is not always clear. Our focus here is on non-efficiency economic values in which political/social values may be embedded.

A classic example of excluding negative economic effects is the Supreme Court's opinion in *National Society of Professional Engineers*. There it was argued (and apparently not disbelieved) that if trade association members were compelled to compete on the basis of price, competition would lead to lower prices. That would lead engineers to reduce the time they could afford to spend on structural design, and this reduced care would lead to catastrophic failures that would be damaging to the economy and to the public safety. The Supreme Court rejected these considerations, saying that the antitrust laws permit only discussion of competitive effects, and the competitive effect of eliminating price from competition was price fixing, which is per se illegal. Assuming (counterfactually, it seems) that the Society's prediction of falling structures was correct, wouldn't this rejection of evidence likely result in major inefficiencies in the economy?

Another Supreme Court example is the opinion in *Superior Court Trial Lawyers Association*. Private criminal defense lawyers requested by the court to represent indigent defendants were markedly underpaid for their work. They announced through their trade association that they would not accept new cases assigned by the court until the city raised their fees. The FTC said this was a per se illegal economic boycott. Affirming, the Supreme Court refused to consider the justifications, including that this was Constitutionally protected political speech—or even to undertake a weighing under a Rule of Reason test. The excluded evidence could have included the inefficiencies of a legal system that purports to provide poor citizens a right to defense but fails to make available sufficient funds to purchase adequate defense services.

Mergers often result in what is known as *delocalization*: the transfer of control of a business from locally based owners or managers to distant or absentee managers. According to Richard Brunell,

A review of the available empirical evidence suggests that the concerns of local community leaders over the loss of a corporate headquarters are justified: delocalization by merger often (but not invariably) involves short- and long-term social costs to the community, including lower civic involvement, philanthropy, employment, and investment.

It is conventional antitrust wisdom in the United States today that social and political values such as decentralization of economic power, environmental health, and public safety are "non-efficiency values" and have no place in antitrust. The localism losses cited by Brunell may indeed reflect a social value, but these costs are primarily economic. Nonetheless, policymakers do not consider the loss of corporate headquarters by merger to be a significant issue.

Taking these kinds of economic values into consideration is said to be, like purely political or purely social values, too difficult, too subjective, too inviting of unpredictable decisions and perhaps even of corruption. (Incidentally, this also raises the question of whether current efficiency analysis is, by comparison, scientific, objective, and prediction-friendly.) At some point these criticisms can undoubtedly

---

54. *Nat'l Soc'y of Prof'l Eng'rs v. United States*, 435 U.S. 679 (1978). This was a price-fixing case arising out of the professional association's code of ethics that restricted the timing for consideration of the fee until after the engineer had actually been selected by the owner. Unlike a merger, where efficiency could be offered as a defense, efficiency would come up in a horizontal collusion case during a Rule of Reason analysis—in the unusual case not treated under the rule of per se illegality—in the context of whether the challenged agreement promotes or suppresses competition.


57. *Id.*

58. But see David W. Barnes, *Nonefficiency Goals in the Antitrust Law of Mergers*, 30 W&M. & MARY L. REV. 787 (1989), who argues that many nonefficiency goals can be incorporated into a broader interpretation of competition. Ben Van Rompuy has shown that the European Union, despite competition guidelines that focus strictly on efficiency, has sometimes incorporated non-efficiency values in its decisions. See infra note 81 and accompanying text.
be true. On the other hand, too narrow a reading of “competition effects” leads to exclusion of many claims that relate to economics and economic efficiency.\(^{59}\) If some types of efficiency are permitted into a case as a defense while others are excluded, there could be a systematic skewing of outcomes that might in itself be considered political, that is, biased in favor of the largest entities attempting to merge or otherwise get their way.

### B. Externalities

A common definition of “externality” is “the direct effect on the well-being of a consumer or the production capability of a firm from the actions of other consumers or firms.”\(^ {60}\) As Carlton and Perloff explain, “An externality occurs when consumers or firms do not bear the full cost (benefit) from the harm (good) their actions do to others.”\(^ {61}\) Thus an externality can be negative or positive. A frequent example of a negative externality is pollution produced by a firm, harming the public but not charged to the firm, hence not being incorporated into the price paid by consumers and consequently leading to overproduction of the polluting product. (Consumers may or may not be affected by the externalized costs in their role as taxpayers or, for example, by personal health problems caused by pollution.) An example of a positive externality is an invention by a firm that is widely copied without compensation to the firm, leading to underproduction of inventions. Statutes, such as patent law, may attempt to rectify the problem.

Because modern antitrust analysis is almost exclusively concerned with price theory and near-term effects on price and output, most externalities are by definition excluded from contemporary antitrust analysis. There is, however, one major exception: the weight given to free-rider arguments in vertical restraint cases. For instance, in the Supreme Court’s Leegin decision\(^ {62}\) reversing the long history of rejecting resale price maintenance as a per se illegal form of price fixing, a key piece of the court’s logic was the concern that discounting retailers would take a “free ride” on the educational services provided by full-price retailers with trained clerks who take time to educate customers. In other words, the majority believed that a retailer who provided a high level of customer service would be creating a positive externality that would accrue to the benefit of the competing discounter whose lower costs/lower prices were made possible by not having to take on the full-service function. In order to eliminate the possibility of free riding, the Court held that it may be reasonable for a manufacturer to set and enforce a retail price high enough to facilitate the retailer’s provision of the functions that the manufacturer deemed important to the proper sale and use of the product. Restrictive behavior by manufacturers or otherwise independent retailers would be allowed so that the externality of free riding could be internalized in the price of the product.

One of the problems in dealing with efficiencies is that they often constitute a moving target. Marketing professor Gregory Gundlach has been studying the growth of multichannel distribution for the American Antitrust Institute.\(^ {63}\) This is the developing pattern of firms reaching out to customers

---

59. See Brunell, supra note 57, at 201–03 and sources cited therein.
60. Carlton & Perloff, supra note 15, at 781.
61. Id. at 82.
through brick and mortar stores, the Internet, and other channels, in a strategically coordinated way. Professor Gundlach challenges the assumptions underlying the free-rider theorizing and observes:

Although antitrust has accepted as fact the theoretical justification of free riding without empirical evidence, the strategy literature documents that (1) in the past companies employed segmented channels that deterred free riding and consumers faced too many costs in attempting to shop across stores and thus engage in the consumer behavior sufficient to raise concerns about free riding; (2) over time companies have developed many alternatives to Retail Price Maintenance that more effectively and efficiently accomplish the purpose of encouraging reseller promotion—the primary goal of curbing free riding; and (3) most recently through the use of multichannel systems and effective multichannel management, firms are actually profiting from consumers shopping across channels in route to a purchase—thus effectively embracing the consumer behavior necessary for “free riding” to occur. In short, the free riding justification for vertical restraints likely was overblown in the past, many alternatives beyond restricting competition exist and are in use today, and the future has companies embracing what was once a key antitrust justification for competitively restrictive behavior.64

Based on the law’s current treatment of Retail Price Maintenance as subject to the Rule of Reason, in which avoidance of the alleged free-rider problem is justification for vertical price fixing, it seems that antitrust law permits some flexibility in picking and choosing when to recognize externalities. On the other hand, while externalities that undermine efficiency are usually excluded from consideration, in many circumstances efficiencies are very much included. This raises a critical question: if antitrust is supposed to contribute to economic welfare, and if internal efficiencies are taken into consideration (i.e., those factors that might increase efficiency and therefore tend to reduce prices or increase output), what are the implications if externalities are excluded? More to the heart of the matter, what if externalities are viewed as inefficiencies ancillary to the drive for efficiency because of the costs they impose on the society at large (or subgroups) when they are not internalized within a transaction?

Example: Fragility of supply chains. Just-in-time delivery has become something of a fad in the world of supply chain management. In effect, it eliminates a redundancy, namely the storage of similar inventory at more than one location at the same time. This is clearly a productive efficiency in that it aids in the reduction of inventory management costs. Indeed, many mergers are allegedly beneficial because of their elimination of all sorts of redundancies, which very often include labor. Unemployment compensation and related costs of lost jobs are externalities of such mergers that are unrecognized in the antitrust analysis, but my point here is that too much redundancy elimination can create fragile supply chains. If, far down the supply chain, there is a bottleneck, say a single supplier of a necessary but small component that is purchased by all the members of an industry, and something disruptive happens to the bottleneck supplier (e.g., a fire or earthquake or tsunami), the ripple effects throughout the industry could include a significant shutdown of supply to millions of dependent manufacturers, distributors, and consumers.65

64. Email from Gregory Gundlach to author (Mar. 7, 2014) (on file with author).
65. BARRY C. LYNN, END OF THE LINE: THE RISE AND COMING FALL OF THE GLOBAL CORPORATION ch. 8 (2005). See Siddhartha Mahanta, New York’s Looming Food Disaster, THE ATLANTIC (Oct. 21, 2013), www.theatlanticcities.com/politics/2013/10/new-works-loomng-food-disaster/7294/ (*As systemic risk expert Charles Perrow has written in books such as The Next Catastrophe, private companies aren’t likely to step up on their own. Market economies are reluctant to bear the costs of redundancy and stockpiling—the incentive to plan for disaster, it seems, isn’t apparent. More alarming: in June, New York City officials released a lengthy report reviewing their response to [Hurricane] Sandy. It received near-universal praise, but failed to account for the revolutionary effects of private supply chain consolidation and just-in-time practices. Instead, it notes that a more comprehensive study of food supply vulnerability is needed, but is contingent on available funding. Governments must set new rules ensuring that there’s always enough food flowing through a resilient*).
The threat of catastrophic breakdown could be reduced by assuring multiple diverse sources or by requiring at-risk companies to hold large reserves, or, at least in theory, by insurance. Any of these potentially price-raising options would constitute the internalization of what is now an externalized cost. But in my example and in real life, there is no insurance. The absence of adequate protection against risk to the industry and to the firm and its various stakeholders should be viewed as an inefficiency caused by the single-minded pursuit of efficiency—the over-reduction of redundancies—by the firms in the industry, accepted by a system that doesn’t take important downsides of consolidation into account.

Example: Harm (and benefit) to communities. Another example might be the merger of two firms in which each had previously been headquartered in a separate city. Both headquarters were important to the host city, for example, as a resource for employment, philanthropy, and community leadership. The merger is premised on (and will likely be financed in part by) the closing down of one headquarters, thereby eliminating redundant management, labor, and overhead expenses.

This is undoubtedly efficient in some sense. But what about the losses that will be suffered by the former headquarters city? Its residents bear the losses, economic, political, and social, without compensation. These costs of the merger are externalized. (At the same time, the positive economic gains to the enlarged headquarters city are also not counted, and there is no reason to think that the losses and gains are offsetting.) The absence of a governmentally imposed compensation fee means that the merger is allowed on the basis of purported efficiency gains, but inefficiencies imposed on the public (and various of its subgroups) are not netted out.

The result is that a policy decision within the antitrust community about what to exclude from antitrust analysis imposes a cost on the public (or some parts of the public) while subsidizing those in the private sector who have a financial interest in merging or, more generally, in merger activity. The decision to exclude consideration of externalities should be viewed as an important developement in politics, defined as “the process that determines the authoritative allocation of values.”

C. Inefficiencies

Thus, some externalities may be caused by the corporate reach for efficiency. In these situations, the market should be considered inefficient because all the costs are not included within the relevant marketing decisions, thus causing the market to operate on incomplete information and misleading signals. There are other kinds of inefficiencies that reflect internal issues rather than uncounted externalities.

distribution network. Because allowing supply chains to break down isn’t just bad public policy—for private companies, failing to invest with an eye towards eventual calamity is just bad business.” A related example is the heavy use of hubs for air traffic. When they work well, the efficiencies are notable. But the fragility of these networks becomes obvious when bad weather in one part of the country sets off a domino effect of delays in other parts of the country.


67. Quoting the eminent political scientist David Easton, who according to the New York Times on February 8, 2014, “provided this definition memorized by a generation of students.” A correction on February 12 clarified that the quote is from Easton and not from the late Robert A. Dahl, as the earlier article had stated. http://www.nytimes.com/2014/02/12/pageoneplus/corrections-february-12-2014.html?_r=0; http://www.nytimes.com/2014/02/08/us/politics/robert-a-dahl-dies-at-98-defined-politics-and-power.html?_r=0.

68. In an environmental book that explores what he calls “the efficiency paradox,” Steve Hallett argues that analyses of efficiency need to include “rebound effects.” For example, when efficiency reforms reduce the costs of a resource, this increases demand, such that ultimately more of the resource is used. Steve Hallett, THE EFFICIENCY TRAP: FINDING A BETTER WAY TO ACHIEVE A SUSTAINABLE ENERGY FUTURE (2013). This could be considered another possible inefficiency of efficiency.
I. X-Inefficiency. In 1966, the economist Harvey Leibenstein published an important article called *Allocative Efficiency vs. X-Efficiency* in the *American Economic Review*. Later, he referred to the same phenomenon as “X-inefficiency.” Leibenstein committed the economic sin, even prior to the Chicago School’s ascension, of “questioning whether market forces could be assumed to ensure allocative efficiency.”

Leibenstein was concerned that the nature of monopoly was being misunderstood by neoclassical economists. Reacting to an analysis by Arnold Harberger which seemed to show that a monopolist’s impact on economic welfare was rather small, Leibenstein insisted that absent competition, firms are unlikely to use their resources efficiently. Adam Smith and Alfred Marshall had noted the importance of competition for efficiency, and many economists had documented the concept of managerial slack. Leibenstein added new information (e.g., anecdotes of previously protected firms that were suddenly confronted with serious competition and suddenly found ample ways to cut expenses) and a kind of theoretical framework, reminding economists of the principal-agent problem, in which both management and labor lack the motivation to maximize firm efficiency. His claim, that deviations from cost minimization are pervasive, presented a rather fundamental challenge to microeconomics, and in fact he found little support from mainstream neoclassical economists.

George Stigler notably attacked Leibenstein for spotty evidence and lack of a formal theory. According to economist Michael Perelman, Stigler’s defense incorporated nonmarketed managerial welfare as a profit-maximizing strategy. One doesn’t hear much of Leibenstein or X-inefficiency today, but economic behavioralism has re-emphasized recognition of the flexibility that exists within a large corporation, or indeed any firm that is not under the immediate threat of extintive competition, to deviate from the maximally efficient posture.

How important is X-inefficiency? To what extent is it found at various levels of market power? In 1990, Scherer and Ross summarized the evidence in this way:

> The evidence is fragmentary, but it points in the same general direction. X-inefficiency exists, and it is more apt to be reduced when competitive pressures are strong than when firms enjoy insulated market positions. What we do not yet know is the magnitude of differences systematically correlated with monopoly power. It seems eminently plausible, however, that X-inefficiencies attributable to monopoly are at least as large as the welfare losses from resource misallocation.

Fifteen years after Scherer and Ross, however, Carlton and Perloff note that many economists reject the idea of X-inefficiency because monopolists, like other companies, want to maximize profits and the only way to do that is to minimize its costs at its chosen output level. This argument assumes the Chicago School model of rational behavior, full information, and profit maximization and ignores

---

73. *Id.*, at 219.
75. CARLTON & PERLOFF, *supra* note 15, at 94. Another economics text of similar vintage states: “Of course, X-inefficiency is inconsistent with the assumption that monopolists maximize profits. However, some economists have argued that the separation of ownership from control in large firms with market power permits the managers to substitute their own objectives for the profit objectives of the owners. Therefore, in such cases, X-inefficiency may arise.” W. Kip Viscusi, ET. AL., *ECONOMICS OF REGULATION AND ANTITRUST* 83 (2d ed. 1995).
possible managerial motivations that can be manifested in high salaries and bonuses, reduced work hours, luxury corporate jets, and so on. In one small concession, Carlton and Perloff note that a monopoly may have more difficulty than a competitive firm in ascertaining what other firms are doing, so it may be more difficult to judge how efficiently it is operating.\textsuperscript{76}

Economist Roger Frantz has reviewed 150 articles on the role of X-inefficiency in the last twenty years, many of which draw on studies of the liberalization of the banking sector across the globe. Several of his conclusions are of note: (1) X-efficiency can be measured; (2) it can be used in merger analysis; (3) X-inefficiency is real and substantial; and (4) U.S. courts have never used an X-efficiency argument in an antitrust case.\textsuperscript{77} This should be a fertile ground for creative antitrust thinking.

2. Diseconomies of scale, scope, and coordination. The literature also concerns itself with diseconomies of scale, scope, and coordination. In other words, as firms grow larger and more complex, after some point they may become less efficient and more difficult to manage. Themes like communications costs, duplication of effort, management opportunism, principal/agent problems, cannibalism, isolation of decision makers, slow response time, and inertia emerge as counterweights to claimed efficiencies of larger scale or scope.

Economies of scale are probably the primary drivers of cost reductions, particularly in industries with high fixed costs. The age-old question is: when are economies of scale exhausted? Once exhausted, additional scale may lead to diseconomies. These could imply inefficiencies because, for example, investing in even greater scale may mean not investing in more efficient alternative expenditures for the firm.

Economies of scope can also be important drivers of cost reductions. The theory is that if a single firm can produce multiple products at lower costs relative to multiple firms each producing individual products, overall costs are lower. The question is, to what extent are scope economies realizable and under what types of multiproduct firms? Moving beyond some tipping point, additional scope may also lead to diseconomies.

Economies of coordination flow from better integration of different levels of production (e.g., specialized boat motors and boats), ensuring quality control, securing sources of inputs, and so on. These vertically separated production costs are difficult to verify. That's why vertical mergers rely foremost on the elimination of double margins (e.g., elimination of the markup in input procurement) as the primary theory when discussing cost efficiencies. Vertical or contract integration can reduce costs up to a point, primarily transaction costs, but after the benefits decline, transaction costs might very well go up. Based on discussion with antitrust economists, it seems fair to observe that economists today are better equipped to analyze efficiencies in horizontal situations than in vertical. "Transaction costs" and "coordination" are concepts that still seem to be difficult to measure. This may help explain why there are no federal vertical merger guidelines and why relatively few vertical mergers are challenged.

Whereas X-inefficiency is associated with the absence of effective competition, diseconomies of scale, scope, and coordination arise from overly large firm size. While, in theory, profit maximizing will deter management from seeking greater scale, scope, or coordination, it is an empirical question whether the diseconomies begin to outweigh the economies at some point in the process. Certainly,

\textsuperscript{76} In theory, capital markets would direct investment away from monopolists that do not maximize efficiency, but investors would have a difficult time determining the extent of X-inefficiency and may be sufficiently pleased by monopoly rents that they do not care about the monopolist's laxness.


\textsuperscript{78} Farrell and Shapiro argue that simple claims of scale economies should be viewed skeptically by the antitrust agencies, but that greater weight should be given "to credible claims of genuine (and merger-specific) efficiencies based upon the close integration of specific hard-to-trade assets owned by the merging parties." Joseph Farrell & Carl Shapiro, \textit{Scale Economies and Synergies in Horizontal Merger Analysis}, 68 ANITRUST L.J. 685, 687 (2001).
experience teaches that empire building can be a powerful motivation for many executives and that forces other than efficiency may provide the chief incentive for various management decisions. Behavioral economics and/or strategic considerations may provide explanations for kinds of decisions that do not appear to be rational in the profit-maximizing sense.

3. The too-big-to-fail problem. The Too-Big-To-Fail debate has exploded since the financial meltdown of 2008. It is to some extent about whether the largest financial institutions have grown so large, so complex, and so interconnected with other parts of the economy that they are in fact too difficult to manage efficiently and without putting the general economy at risk of their failure. 79

In effect, investors view too-big-to-fail banks as less risky than smaller banks, giving them an advantage in raising capital and a competitive advantage vis-à-vis banks that are more realistically subject to failure. Big gets bigger as a result of a size-generated subsidy. This also exemplifies a larger problem in terms of efficiency, namely that whenever government policies, spoken or unspoken, formal or informal, deliver subsidies (or particular disadvantages) to a class of firms, inefficient signals are given that undermine allocative efficiency.

The concept of systemic risk created by “too-big-too-fail” banks has been treated, to some extent, as an externality by subsequent regulatory reforms. The risk is supposed to be internalized to the extent that very large banks have been required to hold additional reserves or equity. Moreover, a new governmental council has been created to monitor and reduce potential systemic risks. The intense political conversation about too-big-too-fail enterprises and what to do about them is likely to continue. One effect has been to delegitimatize discussion, initiated by Louis D. Brandeis in the debates leading up to the creation of the FTC, of the relative value of extremely large organizations, both in terms of the inherent limits on their efficiency and on their impact on a democratic polity.

VI. Reform: Problems of Prediction and Quantification

I have suggested on one hand that antitrust analysis has become increasingly comfortable, at least in theory, with consideration of predicted efficiencies as a justification for what would otherwise be considered anticompetitive situations. One might think that this inevitably raises the question of whether it is appropriate to give defensive weight to efficiencies while ignoring associated inefficiencies that might countervail if considered. Relatively little seems to have been written about this.

While some efficiency claims can be reduced to dollar amounts, the types of external and internal inefficiencies that have been described are notoriously difficult to quantify, much less predict with substantial reliability. For example, the following questions would have to be addressed if both efficiencies and inefficiencies were to be taken into account: (1) What categories of both efficiencies and inefficiencies would be included (cognizable) in the calculation? (2) If various types of efficiency/inefficiency claims are to be taken into account, how far into the future should they be postulated? (3) At what discount rate should they be brought back to present value? (4) Since predictions represent probabilities rather than certainties, what percentage should be applied to any particular probability? And, (5) What techniques of forecasting should be utilized and by whom?

Prediction of inefficiencies creates real challenges, but they are not necessarily different from problems associated with the claims that a particular industry structure, merger, behavior, or antitrust remedy will produce specific efficiencies. The Merger Guidelines require that cognizable efficiencies be “verifiable.” 80 Whatever this means in practice, a similar test ought to be imposed on any

countervailing inefficiencies. The problem is, however, that predictions can only be verified in the future. How would you verify that there is a fifty percent chance of snow tomorrow? You might verify that your favorite weather forecaster says, or, better, the consensus of expert weather forecasters say that this is the best prediction. You might verify that each step in the experts’ line of reasoning that leads to the prediction is based on a reasonable assumption, that the probability applied to each step seems reasonable, and that all reasonable scenarios have been worked out and compared. This might allow you to conclude that the prediction is reasonable, but such a conclusion is still quite different from the conclusion obtained by scrutinizing a past event that can be carefully defined and measured objectively. So, what can we say about future-oriented claims of efficiency and inefficiency?

It may be helpful to know that other countries with competition laws face comparable problems. The European Commission (“EC”) has been moving in the direction of the United States by making consumer welfare and efficiency a single-minded objective of the Competition Directorate, the European Union’s central antitrust enforcement authority. Ben Van Rompuy has written a volume comparing the United States and European approaches to efficiency and non-efficiency values in competition law. He is critical of guidelines that were adopted by the EC for Article 101 TFEU (the equivalent of Sherman Act Section One’s prohibition of unreasonable restraints of trade), which seemingly eliminate non-efficiency considerations. He demonstrates that this evolving policy conflicts with the multiple purposes for which European antitrust was intended; that there are at least seven methods by which non-efficiency considerations can enter into European antitrust decisions; but that the result is an obtuseness that reduces the transparency of the system and undermines the role of such non-efficiency values as public health, cultural diversity, and environmental protection.

Van Rompuy suggests that the EC should provide guidance that

(1) carefully circumscribes situations where conflict between competition concerns and non-efficiency values, enshrined in the [European] Treaties, might occur; and (2) suggests that the most appropriate balancing method to resolve this conflict, subject to the principle of proportionality and the need for transparency. The Commission has built up plenty of experience in this regard. It would be foolhardy to cast aside this experience.

Unfortunately, Van Rompuy has to confess that he “does not provide a blueprint for the guidance that is called for.”

Not all foreign jurisdictions are as fixated on efficiency as the United States and the European Commission. “Most other competition law systems,” writes David J. Gerber,

pursue several objectives, not only in the language of their statutes, but also in the decision making of competition authorities and courts. Often economic development is a central goal, but political goals such as dispersion of power and social goals such as increased access to markets are also common. In addition, fairness has been a major goal in many systems, and it is often seen as necessary for attracting political support for competition law.

81. Van Rompuy is a senior researcher at the T.M.C. Asser Instituut (The Hague), professor in competition policy and media regulation at the Free University of Brussels, and a member of the American Antitrust Institute Advisory Board.
82. See VAN ROMPUY, supra note 11.
83. Id. at 400.
84. Id. at 403.
85. Id. at 404.
86. DAVID J. GERBER, GLOBAL COMPETITION: LAW, MARKETS, AND GLOBALIZATION 264 (2010).
There will be room for many doctoral theses to try to make sense of how these various “public interest” provisions will be integrated into antitrust analysis. Thus far, I am not aware of much guidance on the handling of inefficiencies that we can take from other countries’ experience.

It seems to me that there are three primary roads that can be taken by the United States: (1) limit efficiency/inefficiency concerns to prosecutorial discretion; (2) craft guidelines and case development along lines that restrict consideration of both efficiencies and inefficiencies to those that are determined to be “cognizable”; or (3) invite parties to present their best evidence on both efficiencies and inefficiencies, and let the authoritative decision maker figure out what makes the most sense.

For all of its imperfections, the goal of static (allocative and productive) efficiency is too essential to economic welfare for antitrust to totally ignore. Dynamic efficiency, which is probably the more important concept in the long run if growth is accepted as a key economic goal, is even more difficult to predict than static efficiency. Nonetheless, it must be a fundamental part of analysis, if only because of the centrality to antitrust analysis of entry barriers and the probability of outsider entry into a market, not to mention that innovation leads to more competition and to choice for the consumer. Thus, total elimination of efficiency from the antitrust dialogue is a non-starter.

Prior to its incorporation into the Merger Guidelines, efficiency was said by enforcers to be taken into account during investigations.  

One possibility is to return to those good old days: leave efficiency/inefficiency to the discretion of the enforcers, but keep it out of formal guides and court decisions. If we do this, my prediction is that prosecutors, pursuing the public interest but imbued with the litigator’s competitive determination to win, will probably not give efficiency the benefit of any doubt if they know they can win a case without efficiency being contested. At least on occasion, this might lead to results that would, on net, be sharply inefficient. Of course, if neoclassical economists continue in the substantial enforcement role they have occupied in the antitrust agencies since the 1980s, they may have the internal clout to impose efficiency-based discretionary decisions. The more imposing drawback, therefore, is the absence of transparency. The public would never know what was truly at stake, how seriously efficiency concerns were scrutinized, or what tipped the decision. Without transparency, future improvements in predictability would not be likely and businesses would have a difficult time in determining which of their possible strategies would be challenged as illegal.

It would be possible in theory to require disclosure by the agencies of whether and how they took efficiencies and inefficiencies into account. This would likely be resisted, however, on grounds of both resource commitment and confidentiality. Explaining in a useful way the bases for a decision about efficiency or inefficiency would likely need disclosure of the most essential kinds of proprietary information, involving not only that of the firm or firms in question but also of others in the industry.

Antitrust policy planners did not come easily to the formal inclusion of efficiencies in the Merger Guidelines. Only “cognizable” efficiencies are considered. Conceivably, certain types of inefficiencies—those most likely to occur within a designated (albeit arbitrary) period of time, of largest absolute impact on the overall economy or on certain consumer groups or strategic business groupings—could be designated as “cognizable,” thereby limiting the scope of inefficiency analysis. Unfortunately, the cognition approach, whether applied to efficiencies or inefficiencies, is necessarily arbitrary, meaning that it may ignore what could well be the most important effects. This can be especially troublesome if only quantifiable categories of information are considered, and if one side

---

87. See Thomas B. Leary, The Essential Stability of Merger Policy in the United States, 70 ANTITRUST L.J. 105, 112 (2002) (claiming that the objection to an exclusive focus on economics “has simply vanished from the mainstream debate over antitrust policy” and that he was not aware of “non-economic” factors playing a part in the final decision of the FTC or Antitrust Division in merger cases in the twenty years preceding his writing). Compare with Robert Pitofsky, Antitrust Modified: Education, Defense, and Other Worthy Enterprises, ANTITRUST, Spring 1995, at 23, 23 (stating that the “reality is that as a matter of prosecutorial discretion ... prosecutors ... will occasionally take social welfare considerations into account”).
of the equation is easier to quantify than the other. Defining what is “cognizable” is most likely going to focus on what can be quantified, and in most cases this will be more feasible for claimed efficiencies than inefficiencies.

The third alternative would seem to be to welcome all evidence, on either side of the balance. The benefit of this kitchen sink approach is that the whole picture of effects would be encompassed. This would certainly—initially, at least—make the analytical process more complex, expensive, and time consuming, without providing predictability. In fact, if everything is relevant and no specific factors are given ascertainable weights, predictability will likely be minimized. If predictability is to be generated over time, it will eventually come from opinions by courts, as certain types of arguments are given credibility or dismissed and legal/economic teams will make increasingly well-informed decisions on what resources are justified to be spent on case preparation. A noteworthy advantage would go to the side with the deepest pockets to pay for the most extensive economic research.

One drawback is the diversity of our judicial resources. Few judges have an economics background, much less antitrust training or experience. When we consider how many years have gone by without Supreme Court guidance on merger policy, one should not be particularly optimistic about the development of antitrust policy through the common law. The common law process could be sped up, however, if the enforcement agencies can create credible, detailed guidelines that might be persuasive to courts. The federal Merger Guidelines have, rather remarkably, displaced Supreme Court intervention in merger law.

On balance, I think the second road, which seems similar to Van Rompuy’s call for guidance from the EC, makes the most sense. Establishing a limited range of both cognizable efficiencies and cognizable inefficiencies, as determined by enforcement agency guidelines and judicial oversight, allows both the inclusion of more relevant information and consideration of administrative constraints.

Because the question of efficiency arises most frequently in mergers and seems most difficult when prediction is required, the initial reform focus might be limited to mergers and further limited to the small subcategory in which a second request is made by either the DOJ or the FTC. Rarely will a merger outside of this category—representing roughly three percent of all mergers subject to premerger notification—be challenged. The other ninety-seven percent are assumed to be the result of efficiency-driven motivations and not of sufficient potential competitive harm to be worth pursuing. With respect to the potentially most dangerous three percent, it would be reasonable to reverse the underlying presumption of efficiency and place the burden on the merging parties to demonstrate that their cognizable efficiencies outweigh not only the competitive harm but also the cognizable inefficiencies that may be caused by the merger.

Defining what will be cognizable in this scheme will of course be challenging. But it will also be challenging to determine how to arrive at a net efficiency finding. In principle, this is a simple matter of subtraction. Take away from the positive efficiencies the negative inefficiencies, and see what’s left. Alas, because this is not likely to be susceptible to reasonable quantification, it will have to be handled by reasoned narrative that applies ranges of probabilities to ranges of effects, with as much clarity as to assumptions as possible. Should this process prove too difficult, the next best solution may be to severely limit the situations in which efficiency itself will be considered, for example, within prosecutorial discretion during the investigation phase and perhaps again during a remedy phase.

Getting the thinking right for mergers will likely result in better analytical tools for other types of antitrust interventions as well, just as the Merger Guidelines’ clarification of market definition has been applied in other situations.

---

88. In my opinion, the assumption of efficiency underlying current enforcement policy is misplaced, but concern should be greatest in the most highly concentrated markets. I would urge that “second requests” should cover about twice as many premerger notifications as has been the case, but that is an argument for another day.
VII. Conclusion
The pendulum has swung too far. Efficiency is a relevant concept, but fraught with difficulty. A narrow interpretation of what is efficient, excluding some factors that may contribute to efficiency and others that may undermine efficiency, has led to a situation that either favors the party that proclaims the future efficiency of the merger it wants to consummate or of the conduct it is already undertaking—or it results in verbal games that at the end of the day, yield the advocates of efficiency very little. To give more consistency to the concept, a broader view is needed—one which weighs what we have called inefficiencies against whatever is to be counted as efficiencies.

But when one considers how difficult it would be to trace all efficiencies and related inefficiencies, how many arbitrary decisions would need to be made, and how difficult it would be to quantify the things being compared—it becomes questionable whether it was worth opening up the Pandora’s Box of efficiency in the first place. If we are to retain the concept, I am suggesting that there be a rethinking of what categories should be considered as cognizable on both the efficiency and the inefficiency side. With respect to mergers, where the concepts are most often at stake and where their execution is most difficult, the mergers that are seriously investigated by the federal enforcement agencies should be subject to a reversed burden of proof, with the merging parties bearing the burden of persuasion that what they are proposing is not only not anticompetitive, but that any claimed cognizable efficiencies will outweigh the cognizable inefficiencies by a substantial margin.

Acknowledgements
The author thanks AAI Research Fellows Amaury S. Sibon and Jose Ziebarth for early research assistance and Eleanor Fox, Gregory Gundlach, Geoffrey Kozen, John Kwoka, and Benjamin Van Rompuy for their comments on a prior draft.

Author’s Note
This article benefits from presentations made and expands on remarks I presented at the American Antitrust Institute’s conferences on efficiency, June 18–19, 2014.

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.