

Article

Why Above-Cost Price Cuts To Drive Out Entrants Are Not Predatory—and the Implications for Defining Costs and Market Power

Einer Elhauge[†]

CONTENTS

I. THE CURRENT STATE OF LEGAL FLUX AND SCHOLARLY DEBATE	690
A. <i>Legal Developments and Ambiguities</i>	690
B. <i>The Inadequacy of Traditional Responses in Either Direction</i>	698
II. DEFINING THE CORRECT COST MEASURE	703
A. <i>The Murky and Divided Nature of the Current Debate over Cost Definitions</i>	704
B. <i>Use Whatever Costs Are Variable During the Period of Predatory Pricing</i>	707

[†] Professor of Law, Harvard Law School. I am grateful for funding by Harvard Law School, the Handler Foundation, and the John M. Olin Center for Law, Economics and Business at Harvard Law School. I have also benefited from comments on earlier drafts by Jonathan Baker, William Baumol, Lucian Bebchuk, Joe Brodley, Allen Ferrell, Bruce Hay, Howell Jackson, Louis Kaplow, Benjamin Klein, Al Klevorick, Michael Levine, Janus Ordovery, Mark Ramseyer, Mark Roe, Hal Scott, Matt Stephenson, Bill Stuntz, and Gregory Werden, and from comments by workshop participants at the American Law and Economics Association annual meetings, the Harvard Law and Economics Workshop, the U.S. Department of Justice, and the U.S. Federal Trade Commission.

C.	<i>Use Variable Costs of the Alleged Predatory Increase in Output That Displaces the Rival, Not of Producing the Predator's Entire Output</i>	711
D.	<i>If Short-Term Pricing Can Deter Long-Term Investments, Then Use Magnitude of Predator Costs for the Sorts of Costs Variable to the Victim, but Look to the Future To Measure Cost Magnitudes</i>	714
E.	<i>If (as Likely) Short-Term Pricing Cannot Deter Long-Term Investments, Then Just Use Those Predator Costs Varied by Its Alleged Predatory Increase in Output</i>	719
F.	<i>Conclusion on the Proper Cost Measure</i>	724
III.	REACTIVE PRICE CUTS TO DRIVE OUT ENTRANTS NEED NOT INDICATE INCUMBENT MARKET POWER—AND THE IMPLICATIONS FOR DEFINING COSTS WHERE COMMON COSTS EXIST	726
A.	<i>Individual Routes in Hub-and-Spoke Systems Cannot Be Assumed To Be Separate Markets</i>	729
B.	<i>Why Competitive Markets May Induce Price Discrimination That Maximizes Output</i>	732
C.	<i>Why Competitive Price Discrimination Will Often Require Reactive Above-Cost Price Cuts</i>	743
IV.	RESTRICTING ABOVE-COST PRICE CUTS HAS ADVERSE EFFECTS EVEN WHEN THE INCUMBENT DOES HAVE MARKET POWER AND IMPLEMENTATION DIFFICULTIES ARE IGNORED	754
A.	<i>Effects on Likelihood and Consequences of Less Efficient Entry</i>	757
1.	<i>Consequences for Less Efficient Entrants Who Would Have Entered Without Any Restriction</i>	758
a.	<i>Why Less Efficient Entrants Often Enter Without Any Restriction on Reactive Above-Cost Price Cuts</i>	758
b.	<i>The Undesirable Consequences</i>	762
2.	<i>Effects for Less Efficient Entrants Whom the Restrictions Encourage To Enter</i>	766
a.	<i>Why Restrictions on Reactive Above-Cost Price Cuts Can Provide Weak Encouragement to Entry by Less Efficient Firms</i>	766
b.	<i>The Effects of (Weakly) Encouraging This Additional Less Efficient Entry</i>	770
B.	<i>Effects on Likelihood and Consequences of Efficient Entry</i>	773
1.	<i>Post-Entry Effects</i>	774
2.	<i>Ex Ante Effects on Creation of More Efficient Entrants</i>	774

3.	<i>The Restrictions Cannot Reasonably Be Construed or Modified To Eliminate Their Adverse Effects on Efficient Entrants</i>	777
C.	<i>Effects for Entrants Who Can Overcome Their Initial Efficiency Disadvantage</i>	778
1.	<i>When Overcoming Incumbent Efficiency Advantage Necessitates Some Deterioration in Incumbent Efficiency</i>	779
2.	<i>When Increased Entrant Efficiency Suffices To Overcome Incumbent Efficiency Advantage</i>	782
a.	<i>Why Such Entrants Would Generally Enter Without Any Restriction on Reactive Above-Cost Price Cuts</i>	782
b.	<i>The Undesirable Consequences</i>	786
3.	<i>Entrants That Share the Incumbent's Declining Cost Curve</i>	786
D.	<i>Ex Ante Effects on Incumbent Incentives</i>	792
1.	<i>The Likelihood and Legality of Encouraging Limit Pricing</i> ..	792
2.	<i>Reduced Incentives To Create Efficient Incumbents</i>	795
E.	<i>Summary of Effects and Assessment of Possible Trade-Offs</i>	795
F.	<i>The Restrictions Cannot Reasonably Be Construed or Modified To Eliminate or Suspend the Market-Power Requirement</i>	795
V.	UNAVOIDABLE IMPLEMENTATION DIFFICULTIES	
	WORSEN THE ABOVE EFFECTS.....	795
	A. <i>When Is the Moment of Entry?</i>	795
	B. <i>Post-Entry Quality Changes</i>	795
	C. <i>Difficulties in Defining the Incumbent Price Floor or Output Ceiling</i>	795
	D. <i>Conclusion on Implementation Difficulties</i>	795
VI.	THE BAUMOL BAN ON IMPERMANENT REACTIVE PRICE CUTS.....	795
	A. <i>Post-Entry Effects</i>	795
	B. <i>Implementation and Incentive Problems</i>	795
	C. <i>Ex Ante Effects</i>	795
VII.	CONCLUSION.....	795

In the early 1990s, antitrust law on both sides of the Atlantic appeared to have reached a consensus that predatory pricing required proof of below-cost prices.¹ But the last few years have witnessed a surprising movement toward prohibiting firms from responding to entry with above-cost price cuts. The European courts got things rolling with a 1996 decision holding it illegal for monopolists to adopt selective above-cost price cuts that sacrificed revenue in order to eliminate entrants.² Then, in 1998, the United States Department of Transportation proposed a regulation banning major incumbent airlines from reacting to entry with above-cost price cuts or capacity increases that resulted in “substantially” lower short-term profits than alternative pricing would have.³ In May 1999, the United States Department of Justice brought the American Airlines litigation based on the similar theory that it was predatory to respond to entry with business practices that (even if above cost) “clearly” sacrificed profits.⁴ This government theory was supported by several expert economists, including the Nobel Prize-winning professor Joseph Stiglitz.⁵ And now, an important new article by Professor Aaron Edlin proposes the even broader rule that, when an entrant charges at least twenty percent below the prevailing price, a monopolist cannot respond with any price cut at all for twelve to eighteen months or until it loses its monopoly.⁶ All of these positions restrict reactive above-cost price cuts (or output increases) even if they result in prices that meet (rather than undercut) the entrant’s price, on the notion that buyers would likely stick with the incumbent unless the entrant can offer a lower price.

The basic concept underlying these new legal developments and proposals is hardly new. Some courts and scholars have long thought reactive above-cost price cuts designed to drive out entrants were predatory,⁷ and the idea was a standard staple of Socratic dialogue in antitrust classes.⁸ The Edlin proposal is the same as Professor Oliver E.

1. See *infra* Section I.A.

2. See *infra* Section I.A.

3. See *infra* Section I.A.

4. See *infra* Section I.A.

5. See *United States v. AMR Corp.*, 140 F. Supp. 2d 1141, 1180, 1191 (D. Kan. 2001).

6. Aaron S. Edlin, *Stopping Above-Cost Predatory Pricing*, 111 YALE L.J. 941, 945-46 (2002).

7. See *Transamerica Computer Co. v. IBM*, 698 F.2d 1377, 1386-88 (9th Cir. 1983); *Int'l Air Indus., Inc. v. Am. Excelsior Co.*, 517 F.2d 714, 724 (5th Cir. 1975); William J. Baumol, *Quasi-Permanence of Price Reductions: A Policy for Prevention of Predatory Pricing*, 89 YALE L.J. 1, 2-3 (1979); F.M. Scherer, *Predatory Pricing and the Sherman Act: A Comment*, 89 HARV. L. REV. 869, 885-90 (1976); Oliver E. Williamson, *Predatory Pricing: A Strategic and Welfare Analysis*, 87 YALE L.J. 284, 290-92 (1977).

8. See, e.g., PHILLIP AREEDA, *ANTITRUST ANALYSIS* 178 (2d ed. 1974) (offering a typical set of Socratic questions to present this (and other) theories of predatory pricing); 3 PHILLIP E. AREEDA & HERBERT HOVENKAMP, *ANTITRUST LAW* ¶ 736c3, at 337 (rev. ed. 1996) (specifically considering and rejecting an Edlin-like ban on any price reduction); 3 PHILLIP AREEDA &

Williamson's famous 1977 proposal, except that it substitutes a ban on incumbents lowering their price for Williamson's ban on incumbents increasing their output for twelve to eighteen months after entry.⁹ Edlin's proposal also has much in common, as he acknowledges, with Professor William J. Baumol's ingenious 1978 idea of permitting reactive price cuts only if they are quasi-permanent.¹⁰ These are legendary economists. The approach of the European Union (EU) and the U.S. Departments, in turn, has roots in various cases and scholarship that defined a predatory price as one that would not maximize profits unless it could destroy or discipline competitors.¹¹ The scholars supporting this approach in writings between 1977 and 1981 included such heavy hitters as Professors Lawrence Sullivan, Paul Joskow, Alvin Klevorick, Janusz Ordover, and Robert Willig.¹²

By the early 1990s, however, this earlier wave of theories seemed safely buried, in an apparent triumph for the Areeda-Turner position that predatory pricing must be below cost. But now they have resurfaced in these modern legal developments, partly because cases and scholars defending the cost-based rule rested on conclusory definitions and contestable claims that above-cost restrictions were less administrable and imposed certain short-term losses in post-entry price competition in return for an uncertain long-term gain if the entrant remained in the market.¹³ This never provided a satisfactory theoretical response to the critics nor addressed practical objections to actual industry behavior under such a rule. Critics were particularly provoked by an apparently serious problem confronting the airline industry.¹⁴ On many routes there is an incumbent airline that dominates business on that route and sells at a price well above

DONALD E. TURNER, ANTITRUST LAW ¶ 714c (1978) (discussing but rejecting the general theory).

9. Williamson, *supra* note 7, at 295-96, 333-36. This is not to say the two agree. Williamson rejected a rule banning incumbents from lowering their prices in response to entry, which Professor Alfred R. Oxenfeldt had proposed in 1976 testimony. *Id.* at 296 n.39, 318-20, 328 & nn.109-10, 338 (referring to this 1970s articulation of the Edlin rule as the "price maintenance" or "price umbrella" rule).

10. Baumol, *supra* note 7, at 4-6; Edlin, *supra* note 6, at 978. Again, this does not mean that the differences are not substantively significant. See *infra* Part VI. Baumol's rule was actually first proposed by Professors Areeda and Turner but rejected by them. Phillip Areeda & Donald F. Turner, *Predatory Pricing and Related Practices Under Section 2 of the Sherman Act*, 88 HARV. L. REV. 697, 708-09 (1975).

11. See *infra* Section I.B.

12. See *infra* Section I.B; sources cited *infra* note 71.

13. See *infra* Sections I.B, IV.E.

14. This was the direct motivation for the Department of Justice and Department of Transportation efforts. See *United States v. AMR Corp.*, 140 F. Supp. 2d 1141, 1145-69 (D. Kan. 2001) (recounting similar examples); Enforcement Policy Regarding Unfair Exclusionary Conduct in the Air Transportation Industry, 63 Fed. Reg. 17,919, 17,920-22 (proposed Apr. 10, 1998). Airlines also form the main examples motivating the Edlin analysis. See Edlin, *supra* note 6, at 942-43, 980-87. This concern with above-cost airline predation even goes back to Professor Baumol. See Baumol, *supra* note 7, at 2.

its costs for that route. Periodically, another airline enters the market at a lower price. The incumbent firm then lowers its price to beat (or match) the entrant. The incumbent never prices below its own costs. But because the entrant has higher costs (or lower quality), it cannot compete at the new price and is driven out of the market. Once the less efficient entrant is safely gone, the incumbent reestablishes the old price.

The concern is that such reactive temporary price cuts not only drive out entrants, but deter similar entry in the future, and thus allow the more efficient incumbent to perpetuate monopoly prices that exceed the price the next most efficient firm would charge. If so, the supposedly certain gains from short-run post-entry price competition never arrive because the entry never occurs, and the long-term loss is experienced with certainty every day. Moreover, although airlines present the concern in particularly stark form, this concern can exist in any industry where incumbent firms are more efficient than potential entrants and exploit their market power (when entrants are not present) to charge prices well above incumbent costs. Indeed, if valid, this concern would overturn a general current skepticism based on the presumption that predatory pricing is rare because it requires the incumbent to sustain losses on a large number of sales.¹⁵ If harmful predation involved profitable above-cost pricing, it would be far more plausible and prevalent.

This is a serious concern that can no longer be suppressed with conclusory labels or contestable claims that ignore the effect on incentives to enter. Unless more seriously addressed, these unanswered concerns about above-cost reactive price cuts will likely continue to influence and expand the development of legal doctrines to deal with those concerns in the United States and Europe, both for antitrust law and regulatory agencies. And such unaddressed concerns will bias conclusions about what counts as a cost whenever a cost-based test is still used. It is thus time to take the idea of restricting above-cost reactive price cuts more seriously. But it is not time to adopt that idea. To the contrary, this Article shows that seriously confronting the idea reveals several heretofore unappreciated flaws in such restrictions.

First, such restrictions will often penalize efficient pricing behavior when incumbents do not even have the market power to restrict output. This is because, in many competitive markets, incumbent firms maximize their ability to incur common costs (and thus create output) by charging high-

15. See, e.g., 3 PHILLIP E. AREEDA & HERBERT HOVENKAMP, *ANTITRUST LAW* ¶ 723b, at 273-74 (2d ed. 2002) (collecting sources and linking them to the argument that the rareness of predatory pricing means courts are more likely to condemn desirable pricing erroneously than condemn predatory pricing correctly). *But see* Patrick Bolton et al., *Predatory Pricing: Strategic Theory and Legal Policy*, 88 *GEO. L.J.* 2239, 2241 (2000) (arguing that modern economic literature contravenes earlier claims that below-cost predatory pricing was irrational).

demand buyers higher prices to get them to cover a disproportionate share of recurring common costs, and charging low-demand customers lower prices that are closer to firms' marginal costs once these common costs are incurred. Competition or low entry barriers will ensure that overall revenue from this output-maximizing price-discrimination schedule does not exceed economic costs. This probably describes airlines, which do not earn positive economic profits but do charge more for a ticket that offers one nonstop flight than for the same flight when bundled with a connecting flight. An entrant who cherry-picks by selling only to the high-value customers at a lower price will thus undercut an output-maximizing price schedule. In order to continue to cover common costs, incumbents will have to react to such entry by lowering their prices to those high-value customers. This reactive above-cost price cut will drive a less efficient entrant out of the market. But this does not mean that the price cut protected incumbent market power and harmed efficiency and consumer welfare. To the contrary, it means that the initial discriminatory pricing schedule never indicated market power, and that allowing the normal competitive process of price cuts to drive out the entrant restores the market to an efficient and output-maximizing state.

Second, even if the incumbent does have market power, and we (heroically) assume away the difficulties of implementing the restrictions, the effects of these restrictions are generally undesirable. This is not because, as commonly supposed, the restrictions exchange a certain short-term loss for an uncertain long-term gain.¹⁶ To the contrary, it turns out to be futile to try to encourage long-term entry with restrictions on reactive above-cost price cuts. Less efficient firms will be driven out when any restriction expires by passage of time or loss of monopoly power and thus will confer no long-term benefit. Firms that are or will become equally efficient will enter and stay in the market even with the prospect of above-cost price cuts and thus will not be encouraged by the restriction. Further, while the restrictions will normally inflict short-term losses, this is not certain. Sometimes the restrictions may weakly encourage additional entry by less efficient firms by marginally prolonging the short-run period during which such entrants can remain in the market, though only if the additional short-run profits provide the marginal profits necessary to make total short-run entrant profits cover entry costs. But even in such cases the effects on prices are mixed because a restriction on reactive price cuts can give incumbents perverse incentives to *raise* post-entry prices to speed the day when the restriction expires. Further, the restrictions will clearly increase prices and harm consumer welfare in the lion's share of cases, when entrants are (or will become) equally efficient or when less efficient

16. See *infra* Sections I.B, IV.E (collecting current sources stating that this is the trade-off).

entrants would have entered even without the restrictions. And in all cases, the restrictions will lower productive efficiency and impose wasteful transition costs. Worse, the restrictions will lessen important incentives to create more efficient entrants and incumbents, which will mean higher costs and lower product quality for society generally.

Third, these adverse effects are worsened by implementation difficulties that are not avoidable but are rather an inevitable consequence of trying to regulate firm pricing, output, and responsiveness to entry. While prior analysis has assumed an unambiguous moment of entry, in fact that moment has many possible definitions. Defining entry as the moment when the entrant actually begins sales would, given the normal lead time for entry, allow the incumbent to make anticipatory price cuts that have the same effect as reactive ones. So would any definition of the moment of entry that does not coincide with the time when entry is first foreseeable. Defining entry as an earlier moment when entry is foreseeable (such as when the entrant first begins to plan for entry) would likely mean any twelve- to eighteen-month restriction would expire by the time the entry starts. One might try to avoid the latter problem by lengthening the period of the price restriction, but the longer the period of restriction, the greater the inefficiencies that will result from uncertainties or inflexibilities in the price floors or output ceilings in the face of changing market conditions. Further, *any* definition of entry that begins before the entrant actually begins sales means that incumbent prices would be artificially elevated during a period when this is not offset by possibly lower entrant prices, thus worsening the likely mix of effects. Another difficulty is that any price floor or output ceiling will provoke inefficient increases in product quality, and any effort to clamp down on that by restricting product enhancements will hamper efficient innovation. Finally, any price floor or output ceiling will create additional inefficiencies because it will either embody an inflexible rule, which will cause inefficiencies in market pricing or output given changing market conditions, or a flexible standard, which will create similar inefficiencies because of application imprecision and uncertainties. The above implementation difficulties cannot be dismissed as mere administrative concerns because their effect is to raise prices, hamper market flexibility, and distort innovation. These harms must thus be added to all the other adverse effects noted above.

In sum, the restrictions will not have any benefit outside the limited case where less efficient entrants face entry costs that are not so low that they would have entered without a restriction and not so high that they cannot recoup with short-run entry, but are in that intermediate range where the marginal prolongation of short-run profits encourages them to engage in hit-and-run short-term entry and exit against an incumbent who was really exercising pre-entry market power. And even in that case, the net effects are

mixed without considering implementation difficulties, and become worse when we do. Further, the restrictions will have clear adverse effects for cases involving any other sort of entrant and also discourage investment and innovation in creating more efficient firms.

These points are all entirely separate from the lively debate about whether *below-cost* predatory pricing should be banned. Many scholars think even *below-cost* pricing should be legal because it inflicts greater losses on the predator than its victims, rarely garners a future recoupment that compensates for losses given time and uncertainty discounts, and can be thwarted by entrant or consumer counterstrategies, all of which make *below-cost* pricing self-detering and too irrational to be credible.¹⁷ Others have reached a different conclusion based mainly on arguments about differential access to capital to cover losses, multimarket reputational effects, imperfect information, or efforts to mislead rivals (or the capital markets that might fund rivals) about predator efficiency or market conditions.¹⁸ This Article takes no position on these disputed issues about the desirability of banning *below-cost* predatory pricing. Rather, I focus on the separate theoretical grounds for rejecting any restriction on *above-cost* predatory pricing.

Underlying all these arguments, however, is the fundamental question of how to define “costs,” an issue now normally resolved by rather atheoretical judgment calls that result in a murky and unsatisfactory doctrine. Any definition of “costs” for a doctrine that bans *below-cost* pricing but not *above-cost* pricing must reflect the rationale for treating the two differently. The rationale for treating *above-cost* pricing as permissible depends, as the above summary makes clear, on the assumption that *above-cost* pricing could not deter or drive out an equally efficient entrant. Likewise, the rationale for banning *below-cost* pricing must be that (if firms did engage in it) such pricing could deter or drive out an equally efficient entrant. It thus makes sense that, if one is going to have a doctrine against *below-cost* predatory pricing, “costs” should be defined in whatever way satisfies the condition that an *above-cost* price could not deter or drive out an equally efficient firm. This test has important implications for which cost measure to use. In particular, it clarifies several longstanding problems in defining the relevant costs for predatory pricing, including what to do when industries have near-zero marginal costs, when equally efficient firms have differing variable costs, when all firms in declining industries have

17. See Frank H. Easterbrook, *Predatory Strategies and Counterstrategies*, 48 U. CHI. L. REV. 263, 269-304, 333-37 (1981); Janusz A. Ordover, *Predatory Pricing*, in 3 THE NEW PALGRAVE DICTIONARY OF ECONOMICS AND THE LAW 77, 79 (Peter Newman ed., 1998) (collecting sources).

18. See Bolton et al., *supra* note 15, at 2247-49, 2285-330 (synthesizing the recent literature); Ordover, *supra* note 17, at 79-80.

marginal costs below their variable costs, and when an alleged predator strategically times low prices after it has made capital investments (and thus has low variable costs) but the rival is deciding whether to do the same. In this way, our inquiry into why above-cost prices are *not* predatory will reveal something important about the nature of what *is* predatory.

I. THE CURRENT STATE OF LEGAL FLUX AND SCHOLARLY DEBATE

Cost-based tests of predatory pricing have been changed or challenged in recent years in ways that suggest legal developments have been, and will continue to be, influenced by the underlying debate on above-cost predatory pricing. After detailing these legal developments, I explain why the largely conclusory points offered by both sides in the current legal and scholarly debate cannot resolve the issue in either direction, which instead requires the more in-depth analysis that the balance of this Article takes up.

A. *Legal Developments and Ambiguities*

In the early 1990s, the law on predatory pricing appeared relatively settled. In 1991, the European Court of Justice held in *AKZO* that when a firm with dominant market power prices below average variable costs, those prices are presumed abusive, and that when it prices above average variable costs but below average total costs,¹⁹ its prices are abusive if they are intended to eliminate a competitor.²⁰ This seemed to imply that prices above average total costs could not be abusive even if coupled with such an intent. In 1993, the U.S. Supreme Court in *Brooke* decided that one necessary element of predatory pricing was proof that the defendant priced below incremental costs.²¹ *Brooke* did not resolve which measure of costs

19. A fixed cost is a cost that does not vary with output levels. A variable cost is a cost that varies with output levels. Total costs are the sum of fixed and variable costs. Average variable costs are the sum of variable costs divided by output. Average total costs are the sum of total costs divided by output. Average total cost thus always exceeds average variable cost since it is the sum of average fixed and variable costs. See 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 735b3, at 367; DENNIS W. CARLTON & JEFFREY M. PERLOFF, *MODERN INDUSTRIAL ORGANIZATION* 28-35 (3d ed. 1999).

20. See Case C-62/86, *AKZO Chemie BV v. Commission*, 1991 E.C.R. I-3359 ¶¶ 70-73 (E.C.J.). Just as section 2 of the U.S. Sherman Act makes it illegal to have monopoly power and engage in unilateral exclusionary conduct, EU Treaty 86 makes it illegal to have a dominant position and engage in unilateral abusive conduct. *Id.* ¶¶ 34-75. But U.S. and EU case law sometimes differ in the precise degree of market power necessary to satisfy the first element, and the type of conduct deemed to anticompetitively violate the second element.

21. *Brooke Group Ltd. v. Brown & Williamson Tobacco Corp.*, 509 U.S. 209, 222-24 (1993). In addition to requiring prices below incremental costs, *Brooke* required proof of two other elements whose precise definition varied with the antitrust statute in question: (1) sufficient market power to have the requisite anticompetitive effect in the market where the predatory pricing is occurring, and (2) a sufficient likelihood of recouping the investment in below-cost prices after rivals were eliminated or disciplined. *Id.* at 224-26 (adopting somewhat higher

should be used.²² But there appeared to be a transatlantic consensus that unilaterally set prices had to be below *some* measure of costs to be considered predatory or illegal.

But now, the law on above-cost predatory pricing is in a considerable state of flux. In 1996, the European Court of First Instance in *Compagnie Maritime* sustained a European Commission ruling that it constituted an abuse of a dominant position to adopt a “fighting ships” strategy of responding to entry by making selective price cuts even though the resulting prices were above costs.²³ The Commission relied on three factors: (1) the price cuts were reactive and selective, having been adopted in response to entry and only for those ships whose sailing dates directly competed with the entrant; (2) the reduced prices met (and once beat) the entrant; and (3) the price cuts reduced defendant profits compared to what they would have been with higher prices.²⁴ The Commission got around *AKZO* by saying that, although this practice was not “predatory” pricing, it was nonetheless abusive.²⁵ The Court of First Instance affirmed, ruling that these three objective criteria meant that the reactive above-cost price cuts did not reflect “normal competition” and were thus abusive.²⁶ The court also suggested more broadly that any above-cost price cut (or other conduct) whose “real purpose” was to strengthen a dominant position by eliminating a competitor was illegal, noting internal documents indicating that the defendant’s purpose was “getting rid” of any independent competitors.²⁷

The European Court of Justice affirmed, declining to rule generally on when it was illegal for a dominant firm to make selective above-cost price cuts to meet an entrant, but holding that such price cuts were illegal when the firm had over 90% market share and had the avowed purpose of eliminating

standards of market power and recoupment likelihood under section 2 of the Sherman Act than under the Robinson-Patman Act). The European Court of First Instance has interpreted EU law to reject any requirement to prove a likelihood that the defendant could recoup predatory prices. *See* Case C-333/94 P, *Tetra Pak v. Commission*, 1996 E.C.R. I-5951 ¶¶ 39-44 (E.C.J.) (affirming the interpretation of the European Court of First Instance). It has also rejected the proposition that the dominant position and predatory pricing have to be in the same market, as long as the firm has a dominant position in some market and the leading position in the market where the predatory pricing happened. *Id.* However, the EU Advocate General had opined that EU law should properly be interpreted to require a recoupment test, *see* *Joined Cases C-395/96 P & C-396/96 P, Compagnie Maritime Belge Transps. SA v. Commission*, 2000 E.C.R. I-1365 ¶ 136 (Opinion of Advocate Gen.), and the European Court of Justice has not yet ruled on the issue.

22. *See* 509 U.S. at 222 n.1; *see also infra* text accompanying note 47.

23. *Joined Cases T-24/93, T-25/93, T-26/93 & T-28/93, Compagnie Maritime Belge Transps. SA v. Commission*, 1996 E.C.R. II-1201 ¶¶ 138-153 (Ct. First Instance). This case often goes under the name *Cewal*.

24. *Id.* ¶¶ 139-141.

25. *Id.* ¶¶ 129, 139.

26. *Id.* ¶¶ 144-145, 148, 153.

27. *Id.* ¶¶ 146-148.

the entrant.²⁸ Likewise, in *Irish Sugar*, the European Court of First Instance held that it was illegal for a firm with 88% market share to engage in above-cost price cuts that were selectively adopted at the border in order to deter entry from an importer.²⁹

So, at a minimum, European law now makes it illegal for a firm with a market share near 90% to respond to entry with above-cost price cuts that are selectively limited to the areas where the entrant competes for the purpose of driving that entrant out. Which other above-cost price cuts might be illegal under European law remains unclear. But the cases suggest that European doctrine might ultimately be interpreted to mean that any above-cost price cut made by a monopolist in reaction to entry is illegal if intended to drive out an entrant, and that such an intent can be established not just by subjective evidence but by objective proof that the resulting price failed to maximize the monopolist's short-run profits.

The law regarding above-cost predatory pricing has also been in some flux on the other side of the Atlantic. In 1998, the U.S. Department of Transportation proposed a regulation banning major incumbent airlines in their hub markets from responding to entrants by cutting prices (or expanding capacity) to a level that, although above-cost, resulted in "substantially" lower short-term profits than alternative pricing (or capacity) would have.³⁰ The Department of Transportation limited its proposed regulation to "major" carriers in their "hub markets" based on evidence that prices in those hub markets were higher than prices elsewhere.³¹ The Department assumed that this effectively established a market power to charge supracompetitive prices in those hub markets,³² but did not say it would require a degree of market power sufficient to constitute monopoly power. After receiving comments, the Department of Transportation at the end of the Clinton Administration announced a decision to pursue this strategy by adjudication rather than by regulation.³³

28. Joined Cases C-395/96 P & C-396/96 P, *Compagnie Maritime Belge Transps. SA v. Commission*, 2000 E.C.R. I-1365 ¶¶ 117-120 (E.C.J.). The Court noted that it would be different if the selective price cuts were justified by lower costs on those sailings. *Id.* ¶ 101.

29. Case T-228/97, *Irish Sugar PLC v. Commission*, 1999 E.C.R. II-2969 ¶¶ 173-193 (Ct. First Instance), *aff'd on other grounds*, C-497/99 P, 2001 E.C.R. I-5333 (E.C.J.). The Court also emphasized that the selectivity of the price cuts was not justified by lower costs in those areas, just by the existence of competition the firm wished to deter. *Id.* ¶¶ 173, 188. The Court suggested that there might be an exception to this doctrine if the entrant priced below cost. *Id.* ¶ 185.

30. Enforcement Policy Regarding Unfair Exclusionary Conduct in the Air Transportation Industry, 63 Fed. Reg. 17,919, 17,920 (proposed Apr. 10, 1998).

31. *Id.*

32. *Id.* It is not at all clear such evidence does actually show market power in individual routes. See *infra* Part III.

33. See U.S. Dep't of Transp., Docket No. OST-98-3713-1846, Enforcement Policy Regarding Unfair Exclusionary Conduct in the Air Transportation Industry: Findings and Conclusions on the Economic, Policy, and Legal Issues 4 (Jan. 17, 2001), at http://dmses.dot.gov/docimages/pdf59/121521_web.pdf.

This regulatory proposal illustrated an important point. Even if barred by antitrust law, theories for banning above-cost predatory pricing can influence the myriad of regulatory agencies that have the power to adopt different rules for a particular industry. True, the Bush Department of Transportation itself seems unlikely to pursue such an approach since its new head filed comments opposing the proposed regulation before he took office.³⁴ But no administration is forever, and there remain plenty of other federal or state regulators who might find the proposal more attractive. Thus, the issue would remain important in the United States even if federal antitrust law were settled.

But in fact, federal antitrust law is not so settled. Notwithstanding *Brooke*, the U.S. Department of Justice in May 1999 brought the American Airlines litigation based largely on the same theory as the Department of Transportation regulation.³⁵ Like the Department of Transportation, the Department of Justice's general theory was that it must be predatory for a monopolist of an airline route to respond to entry by expanding capacity or lowering prices in a way that sacrificed short-term profits (even if prices were still above cost) since such a strategy could only be explained by the long-run goal of driving the entrant out of the market.³⁶ In the alternative, and in an effort to stay within *Brooke*, the Department nonetheless offered four possible cost tests. Two of the cost tests used a measure of fully allocated total airline costs that the Department is no longer pressing on appeal.³⁷ The other two cost tests, which are being pressed on appeal along

34. See Norman Y. Mineta, Docket No. OST-1998-3713-814, Comments to the Department of Transportation's Proposed Enforcement Policy Regarding Unfair Exclusionary Conduct in the Air Transportation Industry (July 24, 1999), at http://dmses.dot.gov/docimages/pdf28/38425_web.pdf.

35. See Complaint of United States, *United States v. AMR Corp.*, 140 F. Supp. 2d 1141 (D. Kan. 1999) (No. 99-1180-JTM), available at <http://www.usdoj.gov/atr/cases/f2400/2438.htm>.

36. See Brief for Appellant United States at 25, 29-31, *AMR Corp.* (10th Cir. Jan. 11, 2002) (No. 01-3202) [hereinafter U.S. Appellate Brief], available at <http://www.usdoj.gov/atr/cases/f9800/9814.pdf>; Redacted Memorandum in Support of the Response of the United States in Opposition to American's Motion for Summary Judgment at 15, 17, 19-22, *AMR Corp.* (No. 99-1180-JTM) [hereinafter U.S. Summary Judgment Memo], available at <http://www.usdoj.gov/atr/cases/f7600/7665.pdf>.

37. These two tests (called Test Two and Test Three) measured whether total revenue on the route was less than fully allocated total airline costs for the route either generally (Test Three) or after the allegedly predatory capacity increases (Test Two). See *AMR Corp.*, 140 F. Supp. 2d at 1179, 1203; see also *id.* at 1175-78 (describing fully allocated cost measures). The district court rejected these tests for two reasons. First, they used total costs rather than incremental costs. *Id.* at 1203. Second, they reflected an arbitrary allocation to individual routes of the joint costs incurred by running a hub-and-spoke flight system. *Id.* at 1203-04. Given that the Department defined the market as the individual route, this meant this cost measure included costs incurred in markets other than the one in which the alleged predatory pricing was occurring. See generally *infra* Part III (discussing hub-and-spoke airline economics). The government has not appealed the rejection of these two cost-based tests. See U.S. Appellate Brief, *supra* note 36, at 1-71.

with its general theory that sacrificing profits is predatory,³⁸ are of more interest here.

Test One measured whether profits on the route declined after the capacity was added, concluding that if it did then the incremental cost of the capacity increase exceeded its incremental revenue.³⁹ But this test necessarily takes into account the fact that adding the incremental capacity lowered prices (and thus profits) on the *non*incremental flights. This amounts to requiring a monopolist to equate marginal revenue and costs, which is precisely the sort of calculation that causes economics texts to predict a monopolist will harm consumer welfare by setting a profit-maximizing monopoly price that is *above* marginal cost.⁴⁰ Thus, although framed as a test of the revenue and cost of the incremental added capacity, this test in effect either subtracted forgone profits on the rest of the route from "revenue" or added those forgone profits to "costs"—either of which converted the seeming price-cost test into a profit-maximizing obligation.

For example, suppose an airline earned \$20.6 million on a route that cost \$18 million to operate, and was contemplating adding a flight that would cost \$500,000 to operate, bring in \$1 million in revenue from passengers on that flight, but reduce revenue for the rest of the route by \$600,000 down to \$20 million. Under Test One, the Department would not compare the additional flight's \$1 million in revenue to its \$500,000 in costs. Instead, the Department would condemn the capacity addition as predatory because it reduced profits from \$2.6 million to \$2.5 million. This effectively either (a) subtracts from the flight's \$1 million in revenue the \$600,000 in profits forgone on the rest of the route (resulting in an incremental revenue of \$400,000 that was less than the \$500,000 in that flight's costs), or (b) adds to the \$500,000 in costs the opportunity cost of the forgone profits on the rest of the route (resulting in an incremental cost of \$1.1 million that exceeded the flight's \$1 million in revenue). While I will defer until Part II how one should measure costs, it is vital for analytical clarity to avoid using cost measures that effectively include forgone profits. Otherwise, one cannot keep predatory theories based on a failure to maximize short-term profits analytically distinct from theories based on pricing below costs. The district court rejected this effort to redefine revenue and costs in a way that imposed a duty to maximize

38. U.S. Appellate Brief, *supra* note 36, at 21-22, 25, 29-31, 48-50.

39. *AMR Corp.*, 140 F. Supp. 2d at 1179, 1200; U.S. Appellate Brief, *supra* note 36, at 21-22; U.S. Summary Judgment Memo, *supra* note 36, at 31.

40. See CARLTON & PERLOFF, *supra* note 19, at 87-92; ROBERT PINDYCK & DANIEL RUBINFELD, *MICROECONOMICS* 334-52 (1989).

profits, as have other courts in the past.⁴¹ But, surprisingly, the Bush Administration has appealed the rejection of this cost test.⁴²

The Department's other test compared the revenue earned from the passengers on the added capacity to the incremental costs of the added capacity.⁴³ The district court concluded that this was also a profit-maximization test, citing testimony by Department of Justice expert Professor Stiglitz that this test also embodied a requirement that the defendant not pass up a clearly more profitable alternative.⁴⁴ But while this accurately characterized Test One, this other test condemned a capacity increase only if it was by itself money-losing in the sense that the revenue earned on the new capacity was less than the cost of adding that capacity, thus putting aside any effect the capacity increase might have on prices and profits on the *non*incremental capacity. For example, the hypothetical described in the paragraph above would *not* be predatory under this test because the \$1 million in revenue for the additional capacity was greater than its \$500,000 cost. Thus, this test was not the same as a profit-maximization test; instead, it amounted to avoiding capacity increases the cost of which inefficiently exceeded their revenue. Unfortunately, the government's briefing did not emphasize this distinction, probably because its general theory (and Stiglitz's) was that it was predatory to sacrifice short-run profits in order to drive out a rival. Instead, it defended this test along with Test One on the grounds that it did not require profit-maximization, but only an examination into whether the capacity expansion was clearly less profitable than the alternative of not expanding capacity.⁴⁵ The district court correctly rejected this as mere semantics, holding there was no substantive difference between a claimed duty to choose a "more" profitable alternative and a duty to "maximize" profits, though the court somewhat unfairly failed to acknowledge that the Department position did condemn only business practices that "clearly" did not maximize short-term profits.⁴⁶

41. See, e.g., *Rebel Oil Co. v. Atl. Richfield Co.*, 146 F.3d 1088, 1095 (9th Cir. 1998); *AMR Corp.*, 140 F. Supp. 2d at 1179-80, 1200-02; William J. Baumol, *Predation and the Logic of the Average Variable Cost Test*, 39 J.L. & ECON. 49, 71 n.20 (1996) (collecting cases).

42. See U.S. Appellate Brief, *supra* note 36, at 21-22, 48-50.

43. This test was called Test Four. See *AMR Corp.*, 140 F. Supp. 2d at 1180; U.S. Appellate Brief, *supra* note 36, at 22; U.S. Summary Judgment Memo, *supra* note 36, at 31-32.

44. *AMR Corp.*, 140 F. Supp. 2d at 1180, 1200-03.

45. U.S. Summary Judgment Memo, *supra* note 36, at 31-32. On appeal, the Department has unfortunately continued to lump together Test Four and Test One and defend them based on this unpersuasive argument, rather than distinguishing the two tests on the ground that one requires avoiding capacity increases that fail to maximize profits, whereas the other test only requires avoiding capacity increases that by themselves lose money. See U.S. Appellate Brief, *supra* note 36, at 48-50.

46. See *AMR Corp.*, 140 F. Supp. 2d at 1180, 1202. The district court also rejected this test based on two other grounds. First, it believed that the average variable cost for the route as a whole was the only appropriate measure of costs because the route was the alleged market. *Id.* at

However this case comes out, the temptation to redefine price and cost in a way that effectively prohibits reactive price cuts that sacrifice short-term profits will remain strong for any court or enforcement agency persuaded by the underlying theory that such reactive price cuts are undesirable regardless of the predator's true costs. Even if federal antitrust courts are not willing to go quite so far, theoretical concerns about reactive above-cost price cuts continue to influence U.S. courts as to which cost measure to use under *Brooke*. In particular, federal courts remain divided on whether to retain antitrust review for pricing that is above marginal or variable costs but below average total costs.⁴⁷ There are many other reasons to disagree about which cost measure to use, including which best assures that equally efficient firms will not be excluded in particular cases.⁴⁸ But courts allowing claims above marginal or variable costs, including most notably the Tenth Circuit in the decision on which the American Airlines litigation was based, have also been influenced by the same concerns that underlie the proposals to ban monopolists from using reactive above-cost price cuts to drive out entrants.⁴⁹

Nor have U.S. courts been shy about changing antitrust law in more dramatic ways as theories of antitrust economics develop. The list of antitrust cases overruled as a result of new economic theory is long

1196, 1198-200, 1202-03. For reasons explained below, this is incorrect as a categorical or even presumptive proposition. Normally, the appropriate measure of variable costs (if equally efficient entrants are to be protected) is not the variable cost of producing the predator's entire output, but the variable cost of making the additional predatory output that replaces the output of the alleged victim. See *infra* Section II.C. Still, it may be appropriate to look to the route as a whole if adding the incremental capacity has positive or negative externalities for the rest of the route—such as when adding flights to complete a schedule of hourly flights offers a collective convenience and flexibility that increases demand for all the flights. The district court suggested that such positive externalities might have existed, though in a way that suggested no proof had been introduced on the issue. See *AMR Corp.*, 140 F. Supp. 2d at 1176. Second, the district court objected that the incremental costs of adding capacity were measured by comparing costs with the additional capacity to what the costs were *before* adding the capacity, rather than comparing them to what the costs would have been during the period of predation without the additional capacity. *Id.* at 1202. The court noted that increases in fuel or labor costs over time could undermine a prepredation baseline. *Id.* This seems correct, though prepredation costs should be an accurate and convenient baseline absent any evidence of an exogenous increase in costs during this time, and the Department on appeal has noted that it introduced evidence, ignored by the district court, that its expert in fact did account for the possibility of such exogenous increases. See U.S. Appellate Brief, *supra* note 36, at 52.

47. 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 723d2, at 278-79, ¶ 724c3, at 289-92, ¶ 737b-c, at 394-402, ¶ 739, at 413, ¶ 741a-c, at 441-45, ¶ 741e, at 449-53, ¶ 741f, at 456-57 (collecting the surprisingly diverse appellate authorities).

48. See *infra* Part II.

49. See, e.g., *Instructional Sys. Dev. Corp. v. Aetna Cas. & Sur. Co.*, 817 F.2d 639, 648-49 (10th Cir. 1987) (condemning prices above variable costs but below total costs because the defendant dropped prices sharply when faced with a rival and then raised prices back to high levels once the rival exited, and in doing so was knowingly sacrificing short-term profits to drive out its rivals).

indeed.⁵⁰ Here that possibility is enhanced because many regard *Brooke's* statement requiring below-cost pricing as dicta.⁵¹ Thus, the existing *Brooke* rule might well be changed if federal antitrust courts come to accept the economic critique. State antitrust courts are also not bound by *Brooke* and are thus free to adopt different interpretations of state antitrust law. And statutory amendment is always possible if Congress or state legislatures become convinced of the merits of proposals to ban above-cost predatory pricing.

In any event, we have long since passed the time when only U.S. law matters in antitrust. With European unification, its markets are often as important as U.S. markets. Further, markets are increasingly globalizing, and the United States and European Union effectively have concurrent antitrust jurisdiction over global markets. This means the EU position on antitrust issues is not just relevant but generally matters more because, as the more aggressive antitrust enforcement agency, it effectively defines the line between legality and illegality in global markets.⁵² If U.S. antitrust law

50. See, e.g., *State Oil v. Khan*, 522 U.S. 3 (1997) (overruling the per se rule against vertical maximum price-fixing announced in a prior Supreme Court case); *Copperweld Corp. v. Independence Tube Corp.*, 467 U.S. 752 (1984) (overruling the doctrine that a corporation could conspire with a wholly owned subsidiary); *Continental T.V., Inc. v. GTE Sylvania Inc.*, 433 U.S. 36 (1977) (overruling the per se rule against vertical nonprice restraints announced in a prior Supreme Court case); *U.S. Steel Corp. v. Fortner Enters.*, 429 U.S. 610 (1977) (holding that the per se rule against tying required independent proof of tying market power, even though prior cases had not required such proof). Even Justice Scalia has written an opinion for the Court agreeing that, despite his own penchant for textual interpretations and the supposed super-strong presumption against overturning statutory precedent, courts are free to develop and change federal antitrust law in a common-law fashion. See *Bus. Elecs. Corp. v. Sharp Elecs. Corp.*, 485 U.S. 717, 731-33 (1988).

51. Because the actual ground for the decision in *Brooke* was that the plaintiff failed to establish likely recoupment, prominent scholars have characterized as dicta its statement requiring below-cost pricing. 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 723d2, at 280, ¶ 724, at 284, ¶ 724c3, at 289, ¶ 735a, at 365, ¶ 737a, at 393-94, ¶ 738a, at 405. But any holding requiring recoupment implicitly requires pricing that incurs some sort of loss, otherwise there is nothing to recoup, as Edlin acknowledges. See Edlin, *supra* note 6, at 942 n.5. Although this forecloses Edlin's own approach, requiring a likelihood of recoupment does not (as Edlin supposes, *id.*) necessarily foreclose all bans on above-cost predatory pricing. In particular, it would not necessarily foreclose the Department of Justice position banning only reactive above-cost price cuts that sacrifice short-term profits, a "loss" that could be said to be "recouped" after the entrant exits. See 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 726c, at 305 (stating that recoupment can be of forgone profits). This position might find obstacles, given Court language requiring a likelihood of the defendant "recouping its investment in below-cost prices," and interpreting its past cases to hold that lowering prices to an above-cost level cannot inflict antitrust injury. *Brooke Group Ltd. v. Brown & Williamson Tobacco Corp.*, 509 U.S. 209, 224 (1993). Nonetheless, one could imagine the argument that, strictly speaking, this phrasing and interpretation were also dicta, and that the narrow holding was to require only proof of some recoupment. My point is not to resolve that issue here, but only to observe that these arguments about *Brooke's* requirement of below-cost pricing possibly being dicta marginally increase the likelihood of a change in law (or could serve as the pretext for one) if such a change were deemed desirable as a matter of antitrust policy.

52. See, e.g., Commission Decision of 03/07/2001 Declaring a Concentration To Be Incompatible with the Common Market and the EEA Agreement: Case No. COMP/M.2220—

does not prohibit above-cost predatory pricing and EU law does, then on global markets it is the European doctrine that trumps. There is thus considerable practical import both in the United States and the European Union in dispelling transatlantic economic theories about above-cost predatory pricing that might influence the development of legal doctrine by the more aggressive courts or regulators of either place.⁵³

B. *The Inadequacy of Traditional Responses in Either Direction*

Why has the early 1990s case law that seemingly established the cost-based rule proven so vulnerable? Probably because the underlying concerns about above-cost predatory pricing have never been satisfactorily addressed. One unfortunate tendency has been to declare victory by definition—asserting that a “predatory” price must be below cost or that

General Electric/Honeywell, at http://europa.eu.int/comm/competition/mergers/cases/decisions/m2220_en.pdf (prohibiting a merger approved by the U.S. Department of Justice).

53. However, some recent cases suggest the importance of emphasizing that the issue whether and when a straight price that is above cost should be illegal must be distinguished from the situation when a seller *conditions* an above-cost discount on the buyer's taking all or a high percentage of its purchases from the seller. Two recent decisions applying *Brooke* to the latter contain language indicating they may have mistakenly confused the issues. See *LePage's Inc. v. 3M*, Nos. 00-1368 & 00-1473, 2002 WL 46961, at *9-11, 14 (3d Cir. Jan. 14, 2002), *vacated for reh'g en banc*, 2002 U.S. App. LEXIS 12419 (3d Cir. Feb. 25, 2002); *Concord Boat v. Brunswick Corp.*, 207 F.3d 1039, 1061-62 (8th Cir. 2000). But in fact the condition means that the latter, while not constituting predatory pricing, can amount to *de facto* exclusive dealing under both U.S. law, see 3A AREEDA & HOVENKAMP, *supra* note 15, ¶ 768b2, at 148; 11 HERBERT HOVENKAMP, ANTITRUST LAW ¶ 1807a-b, at 115-18 (1998), and European law, see Case 322/81, *NV Nederlandsche Banden-Industrie Michelin v. Commission*, 1983 E.C.R. 3461 ¶¶ 75-86 (E.C.J.); Case 85/76, *Hoffman-La Roche & Co. v. Commission*, 1979 E.C.R. 461 ¶¶ 89-91 (E.C.J.); Commission Decision 2000/74/EC of 14 July 1999 Relating to a Proceeding Under Article 82 of the EC Treaty (TV/D-2/34.780—Virgin/British Airways), 2000 O.J. (L 30) 1, 20-21. Other language in *LePage's* and *Concord Boat* indicates that the courts recognized this doctrine but mistakenly assumed that a discount was not conditioned when a higher discount *amount* depended on the buyer purchasing a high percentage from the defendant or when the buyer voluntarily agreed to accept the discount. *LePage's*, 2002 WL 46961, at *12; *Concord Boat*, 207 F.3d at 1044-45, 1059-60, 1063-64. Still, there is other language indicating that the actual holdings of the cases can be limited to the proposition that the claims of *de facto* exclusive dealing were not supported by sufficient proof that the discounts produced substantial market-share foreclosure. See *LePage's*, 2002 WL 46961, at *12-13; *Concord Boat*, 207 F.3d at 1059-60.

If loyalty rebates were never illegal unless the resulting price were below cost, then any firm could immunize its exclusive-dealing agreements from antitrust scrutiny by the simple expedient of inflating the price and then offering a rebate conditioned on exclusivity. A key reason to treat such loyalty rebates differently is that, by foreclosing the market share rivals need to reach the minimum efficient scale, loyalty rebates can raise rivals' costs or exclude them from the market altogether. See generally Thomas G. Krattenmaker & Stephen C. Salop, *Anticompetitive Exclusion: Raising Rivals' Costs To Achieve Power over Price*, 96 YALE L.J. 209 (1986) (describing how foreclosure can raise rivals' costs). When they do so, they exclude rivals not by virtue of advantages they earned by improving their own efficiency, but by reducing their rivals' efficiency. Rewarding the former is socially desirable. Rewarding the latter is not.

low above-cost prices involve “competition on the merits.”⁵⁴ But these are mere formalistic labels that do not answer the substantive question concerning what the law should define as “predatory” pricing or “competition on the merits.” Indeed, the European Commission had a very similar test—whether reactive above-cost price cuts that intended to eliminate rivals involved “normal competition”—and simply drew the opposite formalistic conclusion that they do not.⁵⁵ Assertions about such formalistic labels in either direction do not really aid the inquiry.

Another unfortunate tendency has been to dismiss bans on above-cost predatory pricing with the observation that they protect only “higher cost” or “less efficient” firms.⁵⁶ This observation is important, but by itself does not dictate any conclusion about the social desirability of keeping those less efficient firms in the market, and thus should not be permitted to end the analysis by epithet. Even less efficient firms play a useful role in constraining the prices that more efficient firms can charge.⁵⁷ If proponents are right that restricting reactive above-cost price cuts would increase entry, lower incumbent prices, and enhance consumer welfare,⁵⁸ then keeping less efficient firms in the market may be desirable,⁵⁹ and courts could thus redefine “predatory pricing” to cover (and “competition on the merits” to exclude) any undesirable reactive above-cost price cuts.

The more substantive response has traditionally been to concede that restricting above-cost price cuts often does have long-term benefits on entry and pricing, but to observe that (1) they would raise short-term prices (and lower output) following entry, and (2) it is administratively difficult to sort out when the long-run benefits outweigh the short-run costs.⁶⁰ But point one

54. See, e.g., *Brooke*, 509 U.S. at 223; 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 723a, at 272; 3 AREEDA & TURNER, *supra* note 8, ¶ 714c, at 161; Areeda & Turner, *supra* note 10, at 706-07, 711.

55. Joined Cases T-24/93, T-25/93, T-26/93 & T-28/93, *Compagnie Maritime Belge Transps. SA v. Commission*, 1996 E.C.R. II-1201 ¶ 130 (Ct. First Instance).

56. See, e.g., *Brooke*, 509 U.S. at 223; 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 736b1, at 377, ¶ 736c3, at 384; 3 AREEDA & TURNER, *supra* note 8, ¶ 714c, at 161, 163; RICHARD POSNER, *ANTITRUST LAW: AN ECONOMIC PERSPECTIVE* 188, 193 (1976); Areeda & Turner, *supra* note 10, at 711; Phillip Areeda & Donald F. Turner, *Williamson on Predatory Pricing*, 87 YALE L.J. 1337, 1339, 1342 (1978).

57. In fact, in every market there is some firm that is more efficient than the others. Workable competition is still valuable in such markets. Indeed, even when one firm is so much more efficient that it can be said to be dominant, the existence of the less efficient firms constrains the pricing of the most efficient firm. See, e.g., CARLTON & PERLOFF, *supra* note 19, at 107-18; W. KIP VISCUSI ET AL., *ECONOMICS OF REGULATION AND ANTITRUST* 164-66 (2d ed. 1995).

58. See Edlin, *supra* note 6, at 945-49.

59. Even if prices were lowered, there is the additional question whether this benefit to consumer welfare offsets the loss of productive efficiency that results from transferring market share to a less efficient producer. See *infra* Part IV.

60. *Brooke*, 509 U.S. at 223-24; *Barry Wright Corp. v. IIT Grinnell Corp.*, 724 F.2d 227, 231-35 (1st Cir. 1984); PHILLIP AREEDA, *ANTITRUST ANALYSIS* 196-97 (3d ed. 1981); 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 723d2, at 277, ¶ 735a, at 364-65, ¶ 736a, at 373-75, ¶ 736b-c, at

is hardly a satisfying riposte to the claim that the posited adverse short-run effects never materialize, or are outweighed by long-run benefits, because the restriction on reactive above-cost price cuts would encourage entry (or lower everyday incumbent prices) that otherwise never would have occurred.⁶¹ And point two lends itself to the critique that price-cost comparisons are themselves difficult to administer, and to efforts to make the restrictions more administrable by banning all reactive price cuts or output expansions (like Edlin or Williamson) or at least those that clearly or substantially sacrifice short-term profits (like the Departments).⁶²

For example, the leading antitrust treatise notes no particular administrability problem with an Edlin-like ban on any price reduction, but dismisses it with the simple observation that it would lower the incumbent's post-entry output.⁶³ Why this objection should be a show-stopper is never explained, which seems odd since one of the authors had previously observed that such a price-maintenance rule increased pre-entry output.⁶⁴ This treatise also considers a price floor at the short-term profit-maximizing level (like the one developing in the European Union and proposed by the U.S. Departments), but dismisses it purely on grounds that it is inadministrable.⁶⁵

By the same token, the debate is also not resolved in the other direction by asserting that reactive above-cost price cuts must be illegal because they fit the test of being designed to maintain monopoly power by excluding rivals. As we saw, some language in the European case law seems to embrace this argument.⁶⁶ Likewise, in the United States, proponents have argued that reactive above-cost pricing must be illegal because it fits the basic *Grinnell* test of being designed to exclude rivals and maintain monopoly power.⁶⁷ *Grinnell* stated:

The offense of monopoly under § 2 of the Sherman Act has two elements: (1) the possession of monopoly power in the relevant market and (2) the willful acquisition or maintenance of that power as distinguished from growth or development as a consequence of a superior product, business acumen, or historic accident.⁶⁸

379-93; 3 AREEDA & TURNER, *supra* note 8, ¶ 714c, ¶ 715a, at 166-68; Areeda & Turner, *supra* note 10, at 708-09; Areeda & Turner, *supra* note 56, at 1339.

61. Edlin, *supra* note 6, at 945, 956, 977.

62. *But see infra* Part V (showing why such efforts fail).

63. 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 736c3, at 383.

64. *See* Areeda & Turner, *supra* note 56, at 1340-43.

65. 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 736c2, at 381-82.

66. *See supra* Section I.A.

67. *See* U.S. Summary Judgment Memo, *supra* note 36, at 14-15; Edlin, *supra* note 6, at 965.

68. *United States v. Grinnell Corp.*, 384 U.S. 563, 570-71 (1966).

The second element is often rephrased as requiring “exclusionary” conduct, which is conduct that tends to exclude rivals other than “competition on the merits.”⁶⁹

But, as antitrust scholars have long understood, the problem with the *Grinnell* test is that it is either wrong or conclusory. Often a firm deliberately tries to exclude rivals and acquire or maintain monopoly power with superior products, business acumen, or other conduct that could be considered competition on the merits. The two are not mutually exclusive concepts, as *Grinnell*’s “as distinguished from” language wrongly suggests. In practice, this tension is resolved by court decisions labeling particular conduct that excludes rivals and enhances monopoly power as being either “predatory” and “anticompetitive” on the one hand, or “business acumen” and “competition on the merits” on the other. But without some underlying normative theory to explain when to apply which label, such case law would merely be conclusory.

Nor is the matter settled, as the European Commission and U.S. Departments apparently thought, by evidence that the defendant has sacrificed short-run profits and is thus engaging in behavior that could only be profitable if it had the long-term aim of acquiring monopoly power and earning monopoly returns.⁷⁰ True, such a definition of “predation” has long been advanced by many courts and a long line of distinguished antitrust scholars.⁷¹ But the problem is that this definition would apply equally to all

69. *Aspen Skiing Co. v. Aspen Highlands Skiing Corp.*, 472 U.S. 585, 605 n.32 (1985).

70. *See supra* Section I.A.

71. *Neumann v. Reinforced Earth Co.*, 786 F.2d 424, 427 (D.C. Cir. 1986) (Bork, J.) (“[P]redation involves aggression against business rivals through the use of business practices that would not be considered profit maximizing except for the expectation that (1) actual rivals will be driven from the market, or the entry of potential rivals blocked or delayed, so that the predator will gain or retain a market share sufficient to command monopoly profits, or (2) rivals will be chastened sufficiently to abandon competitive behavior the predator finds threatening to its realization of monopoly profits.”); *Transamerica Computer Co. v. IBM*, 698 F.2d 1377, 1386-88 (9th Cir. 1983) (holding that prices above average total costs can be predatory if “the anticipated benefits of defendant’s price depended on its tendency to discipline or eliminate competition and thereby enhance the firm’s long-term ability to reap the benefits of monopoly power” (quoting *William Inglis & Sons Baking Co. v. ITT Cont’l Baking Co.*, 668 F.2d 1014, 1035 (9th Cir. 1981))); *Janich Bros. Inc. v. Am. Distilling Co.*, 570 F.2d 848, 856 (9th Cir. 1977) (“Pricing is predatory only where the firm foregoes short-term profits in order to develop a market position such that the firm can later raise prices and recoup lost profits.”); *Int’l Air Indus. Inc. v. Am. Excelsior Co.*, 517 F.2d 714, 724 (5th Cir. 1975) (holding that pricing above average variable cost can be predatory if “the competitor is charging a price below its short-run, profit-maximizing price and barriers to entry are great enough to enable the [defendant] to reap the benefits of predation before new entry is possible”); 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 737b, at 396 n.12 (collecting other cases quoting similar tests); LAWRENCE ANTHONY SULLIVAN, HANDBOOK OF THE LAW OF ANTITRUST 113 (1977) (noting that a characteristic feature of predation is a “price substantially below the profit maximizing . . . price,” which thus “makes sense if, but only if, it is seen as a means of driving out or controlling competitors”); Bolton et al., *supra* note 15, at 2242-43 (adopting the same definition and collecting other sources); Janusz A. Ordover & Robert D. Willig, *An Economic Definition of Predation: Pricing and Product Innovation*, 91 YALE L.J. 8, 9-10, 15-16 (1981) (same). According to Joskow and Klevorick:

sorts of desirable conduct. It would apply to any monopolist that does not fully exploit its monopoly power because a failure to charge the full profit-maximizing monopoly price could only be explained by a desire to discourage further entrants.⁷² This would amount to a legal duty to engage in monopoly pricing. Worse, this definition would apply to any firm that invests research and development funds to invent a new innovative product that will allow it to drive out rivals and earn monopoly rents.⁷³ It would also apply to any firm that sacrifices short-term profits by investing in building new facilities, training personnel, or making organizational or distributional changes in order to improve costs or quality and drive out rivals.⁷⁴ Sacrificing short-term profits to build a better or cheaper mousetrap or organization is socially desirable, even though the monopolist is motivated not by any social benefits but by the prospect that eliminating rivals will allow it to reap long-term monopoly profits. Indeed, the prospect of those long-term monopoly profits is desirable precisely because it encourages such efforts.

The proper question thus cannot be whether the defendant sacrificed short-run profits or intended to exclude rivals or gain a monopoly. It is whether the *means* it chose to do so are undesirable in a way antitrust law can regulate without having unduly negative effects on other desirable conduct. And that requires an assessment of the desirability of the consequences of adopting any restriction on reactive above-cost price cuts. It is to that task that I turn next.

Predatory pricing behavior involves a reduction of price in the short run so as to drive competing firms out of the market or to discourage entry of new firms in an effort to gain larger profits via higher prices in the long run than would have been earned if the price reduction had not occurred.

Paul L. Joskow & Alvin K. Klevorick, *A Framework for Analyzing Predatory Pricing Policy*, 89 *YALE L.J.* 213, 219-20 (1979).

72. See *infra* Section IV.D (discussing why limit pricing would violate this proposed test).

73. Anticipating this implication, Professors Ordovery and Willig would actually extend their prohibition to condemn as "predatory" any product innovations whose profitability depends on their ability to drive rivals out of the market. See Ordovery & Willig, *supra* note 71, at 22-30. But this ignores the fact that innovations create long-term positive externalities for society (by lowering cost curves or increasing product value) that matter much more than any short-term loss of allocative efficiency, and that spur a dynamic response of innovation by other firms and entrants that can trump the first innovation. See *infra* Part IV. Our intellectual property laws thus correctly adopt the different premise that it is socially desirable to reward innovations with a right to exclude rivals from their fruits. Further, Ordovery and Willig's test would sometimes prohibit innovation because it sacrifices profits earned on the innovator's older products even if those profits were supracompetitive. See Ordovery & Willig, *supra* note 71, at 25-26. But such a sacrifice of supracompetitive profits is desirable because it brings the quality-adjusted price of market products closer to their cost. Ordovery and Willig wrongly think that such a profit sacrifice can only have an anticompetitive objective. *Id.* at 26 n.49. This ignores the possibility that incumbents fear rival competition in innovation over time, which would naturally tend to squeeze out the supracompetitive profits on the preexisting good unless the Ordovery-Willig test were adopted.

74. Indeed, Schumpeter would say that all innovative investments require such a sacrifice of short-term profits to reap monopoly gains, and thus necessarily require the possession or prospect of some degree of market power. See *generally infra* note 266.

II. DEFINING THE CORRECT COST MEASURE

The effects one predicts from a restriction on above-cost price cuts obviously depend on what counts as a “cost.” The dominant practice has been to rely on a series of judgment calls about both what sorts of costs generally seem sensible to include and when certain cases seem to suggest that ad hoc exceptions to those cost measures would be desirable. But this can hardly offer a satisfactory or clear resolution without some consistent theory to guide those judgments. Nor is this a question we can answer by assuming that “costs” have some metaphysically correct definition. Rather, we must ask *why* we want to know what costs are in order to know how to define them.

To those in the camp that believes that below-cost pricing to deter entrants or drive out rivals is irrational and thus not credible, the answer is that we don’t want to know because their favored doctrine would permit low prices regardless of costs.⁷⁵ But any doctrine that condemns monopolists who engage in below-cost pricing, which is what we actually have, must rest on the opposite premise that such below-cost pricing sometimes is a rational and credible strategy, otherwise the doctrine is not only unnecessary but harmful.⁷⁶ For purposes of such a doctrine, the key reason to care about the distinction between below-cost and above-cost prices must be that (if implemented) below-cost pricing will inflict losses on an equally efficient entrant or rival that can deter its entry or cause its exit. It is further clear that—no matter what theory one holds—pricing that does *not* inflict losses on entrants or rivals cannot deter or drive them out of markets because they will be better off entering or staying in the market no matter what they believe about whether the pricing will occur or persist.

I will accordingly define “costs” as whatever measure of costs would prevent an incumbent pricing at cost from inflicting losses on an equally efficient entrant or rival that could deter its entry or cause its exit. This definition, it will be shown in the rest of this Article, provides the necessary premise for the arguments that follow about why efforts to restrict above-cost price cuts are socially undesirable. In other words, the ultimate justification for this definition is functional. With this definition, one can derive strong functional reasons for distinguishing above-cost price cuts from below-cost price cuts; without it, those reasons would not follow.⁷⁷

75. See *supra* text accompanying note 17.

76. See *supra* text accompanying note 17.

77. Posner advocates such a benchmark but does not justify it. See POSNER, *supra* note 56, at 188. Baumol analyzes a similar benchmark (although limited to pricing that might drive out rivals as opposed to deterring their entry) as a “legitimate borderline” but acknowledges that his own work indicated such a benchmark might reduce social welfare by allowing firms to drive out marginally less efficient firms. See Baumol, *supra* note 41, at 50, 55-57 & n.12. The analysis in this Article provides the necessary justification. This is the correct benchmark because otherwise

Accordingly, while the analysis below necessarily depends on this initial premise about how costs are defined, an important implication of this analysis will be that this cost definition is correct precisely because it is a necessary premise for justifying the distinction between above-cost and below-cost pricing even when both are designed to drive out entrants.

This understanding of the underlying justification for the proper cost definition also has the helpful feature of permitting one to sort through, in a more precise and analytic manner, what is currently a murky and confused debate on how to measure costs. In particular, this Part shows how it can provide a theoretically coherent methodology for resolving many seeming conundrums currently handled largely by ad hoc exceptions, including how to apply a cost-based predatory pricing test (a) to industries with near-zero marginal costs; (b) to equally efficient firms that have different ratios of fixed to variable costs; (c) in declining industries where all firms have marginal costs that are below their variable costs; and (d) when an alleged predator strategically times low prices after it has made capital investments (and thus has low variable costs), but the rival is deciding whether to do the same. It will be useful to sort through these issues not only for their own sake, but because doing so will address in advance a mistaken premise sometimes invoked in models used by those advocating restrictions on above-cost price cuts: that any cost-based test would necessarily deter entry by barring an entrant from recovering sunk entry costs.⁷⁸

A. *The Murky and Divided Nature of the Current Debate over Cost Definitions*

Scholars have taken a variety of positions about the proper cost measure. The leading proponents of a cost-based test, Professors Phillip Areeda, Herbert Hovenkamp, and Donald Turner, argue in their influential antitrust treatise that short-run marginal cost is the correct measure but recommend using average variable cost as an imperfect but more measurable surrogate.⁷⁹ One clear limitation of this approach is that Areeda and Turner acknowledge that their test might allow pricing below average total costs that destroys or deters an equally efficient entrant.⁸⁰ But

the social welfare grounds detailed below for rejecting restrictions on "above-cost" pricing would not apply.

78. See Edlin, *supra* note 6, at 955-60, 973-78. Although Edlin does not extend his model to equally efficient entrants, others have offered models with a similar assumption about entry costs, and thus concluded that (at least when the incumbent is a monopolist in multiple markets) pricing above variable costs can deter equally efficient entrants. See David Easley et al., *Preying for Time*, 33 J. INDUS. ECON. 445, 447-54, 457 (1985).

79. See AREEDA, *supra* note 60, at 194-95; 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 724c1-2, at 287-89; 3 AREEDA & TURNER, *supra* note 8, ¶ 711d.

80. See 3 AREEDA & TURNER, *supra* note 8, ¶ 714a, at 164-68 & n.7; Areeda & Turner, *supra* note 10, at 711-12. They justify this result on the grounds that the alternative is (1) protecting

reflecting a somewhat inconsistent attitude toward this benchmark, Areeda, Turner, and Hovenkamp would switch to average total costs if they are exceeded by the predator's marginal costs in part because prices above average total costs could not drive out equally efficient rivals.⁸¹ They also take the position that when marginal costs fall below average variable costs, one should go by average variable costs not just as a surrogate but on the merits, because if prices are below average variable costs it would be more efficient for the firm to close operations.⁸² To deal with the problem that average variable cost can be below marginal cost, they would also require a defendant to show that marginal costs were not "substantially" or "significantly" above average variable costs.⁸³ Thus, in the end Areeda, Turner, and Hovenkamp really embrace a three-staged cost test: (1) when below the output that minimizes average variable costs, use average variable costs; (2) when between the outputs that minimize average variable and total costs, use average variable costs unless marginal costs are significantly higher; and (3) when above the output that minimizes average total costs, use average total costs. To complicate matters further, they invoke several exceptions from these tests when they seem likely to lead to adverse results.⁸⁴

Although not a proponent of a cost-based predation test, Professor Baumol has argued that the right measure of costs to prevent predation that could drive out an equally efficient rival must be whatever sorts of costs the

some less efficient entry and (2) incurring short-run (and perhaps long-run) market inefficiency since there must be excess capacity for marginal cost to be below average cost. Williamson has properly criticized them for failing to connect the goals of short-run efficiency and minimizing inefficient entry with any larger social welfare calculus, especially since Areeda and Turner concede alternative tests would have the long-term effect of increasing pre-entry output. See Areeda & Turner, *supra* note 56, at 1339, 1342; Oliver E. Williamson, *Williamson on Predatory Pricing II*, 88 YALE L.J. 1183, 1186-87 (1979). Further, the Areeda-Turner test, as stated, encourages the inefficient pre-entry creation of excess capacity that justifies the short-run price below average cost. See *infra* text accompanying note 117.

81. See 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 739a, at 413, ¶ 739c2, at 418; 3 AREEDA & TURNER, *supra* note 8, ¶ 715b2.

82. AREEDA, *supra* note 60, at 195; 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 740b1; 3 AREEDA & TURNER, *supra* note 8, ¶ 715d, at 175.

83. AREEDA, *supra* note 60, at 195; 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 740c; 3 AREEDA & TURNER, *supra* note 8, ¶ 715d, at 176.

84. Areeda, Turner, and Hovenkamp would provide exceptions to their rule condemning prices below average variable costs (a) when the industry has so much excess capacity that all firms are pricing below their average variable cost or (b) when a defendant builds a plant that turns out to be too costly compared to demand, so that prices do not cover a standard measure of average variable costs that includes use depreciation. See *infra* text accompanying note 105; *infra* note 110. They also provide an exception to their rule permitting prices above average variable costs when fixed costs were incurred just to drive out the rival. See *infra* note 110. They also acknowledge that exceptions would be warranted on theoretical grounds (a) when an industry is expanding and variable or marginal costs are low compared to fixed costs and (b) when the defendant retains inefficient excess capacity on hand in order to be able to respond to entry, but they reject these theoretically sound exceptions on administrative grounds. See *infra* note 100; *infra* text accompanying note 117.

rival could avoid by exiting the market.⁸⁵ He calls these “average avoidable costs,” and notes that they exclude inescapable sunk costs “that cannot be avoided for some limited period of time” but include any unsunk fixed costs that “must be incurred in a lump in order for any output at all to be provided.”⁸⁶ Unfortunately, there is some confusion because different authors use different meanings of “fixed costs.” While Baumol defines fixed costs to exclude sunk costs, Areeda, Turner, and Hovenkamp state that “fixed costs are costs that would continue even if the firm produced no output at all,”⁸⁷ which seems to correspond to Baumol’s notion of sunk costs. Baumol’s notion of fixed costs seems to correspond to what Areeda, Turner, and Hovenkamp might call the marginal or variable cost of producing the first unit of output. Some scholars distinguish between sunk and avoidable fixed costs, but describe both as included within the category of fixed costs.⁸⁸

Other prominent antitrust scholars, including those who are ordinarily conservative about antitrust like Judge Richard Posner, worry that a predator’s price could be above short-run marginal, variable, or avoidable costs, yet well below an equally efficient firm’s long-term cost of staying in business. They thus advocate a cost measure that also includes fixed and sunk capital costs—called variously “long-run marginal costs,” “long-run incremental costs,” or “average total costs”—and would condemn prices below that cost measure either presumptively or (under some versions) when coupled with an intent to exclude rivals.⁸⁹ They too, however, invoke exceptions when this test seems likely to lead to bad results.⁹⁰

The cases in both the United States and Europe have responded to this confusion and disagreement mainly by holding that prices between average variable and total costs might be illegal, but differ on against whom to allocate the presumption, and on the grounds for rebuttal.⁹¹ The result is

85. See Baumol, *supra* note 41, at 57-59.

86. *Id.* at 57 n.13, 58-59.

87. 3 AREEDA & TURNER, *supra* note 8, ¶ 712, at 154; see also 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 735b3, at 366.

88. See, e.g., CARLTON & PERLOFF, *supra* note 19, at 28-29.

89. E.g., POSNER, *supra* note 56, at 189, 191-93 (recommending that a price between short-run and long-run marginal cost be considered predatory if coupled with an intent to exclude rivals); Bolton et al., *supra* note 15, at 2271-82 (recommending that a price above average avoidable cost but below long-run incremental cost give the defendant a burden of production (but not persuasion) on whether the pricing maximized short-run profits or had market-expanding efficiencies); Joskow & Klevorick, *supra* note 71, at 252-54 (recommending that a price between average variable and total costs be presumed predatory unless the predator shows it maximizes short-run profits, which is likely only when an industry has excess capacity); Richard A. Posner, *The Chicago School of Antitrust Analysis*, 127 U. PA. L. REV. 925, 942-44 (1979) (recommending the same for a price between average variable and total costs).

90. See *infra* text accompanying note 110 (describing the declining-industry exception).

91. See Case C-62/86, AKZO Chemie BV v. Commission, 1991 E.C.R. I-3359 ¶¶ 70-73 (E.C.J.) (describing European law); 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 723d2, at 278-

that if you are a monopolist or victim and prices are in between these cost measures, you do not really know where you stand.

In short, the current debate is something of a mess. But with our functional criteria in place, we can add some clarity. Much of the problem is that there is little discussion about the actual source of disagreement. The current debate is framed as being about *which costs* to include, when in fact the real debate is about *which time period, output, and firm* to consider in deciding how to categorize a cost. All costs are variable or avoidable in the sufficiently long run.⁹² The fixed costs (like overhead) necessary to make any output this year need not be incurred next year. Generally, even sunk costs are inescapable only for a time. The big expensive plant will eventually wear out and thus require a decision about whether or not to incur the cost of its replacement.⁹³ Even land costs are not inescapable in the long run. Although the land does not wear out, the plant on it does, so that continuing to use the land for present purposes incurs the opportunity cost of not selling the land for its market value. There is thus no cost that is inherently variable, avoidable, fixed, or sunk. It all depends on which time period one uses, whether that period looks backward or forward, and whose output and ability to vary or avoid costs during that period matters. But there has not been much explicit debate about these points, leaving the current analyses murky. We can be more explicit on all these points by considering more directly and systematically the extent to which they bear on the ability of an incumbent pricing at cost to impose losses that could deter or drive out an equally efficient entrant or rival.

B. *Use Whatever Costs Are Variable During the Period of Predatory Pricing*

Professors Areeda, Hovenkamp, and Baumol all state that the correct time period for judging whether costs are variable or avoidable is the time period of the alleged predatory pricing.⁹⁴ But Areeda and Hovenkamp provide no justification for this standard, which they in fact abandon in favor of a blanket assumption of “middle-run” variability,⁹⁵ and the choice requires much more justification and elaboration than Baumol gives.

79, ¶ 724c3, at 289-92, ¶ 737b-c, at 394-402, ¶ 739, at 413, ¶ 741a-c, at 441-45, ¶ 741e, at 449-53, ¶ 741f, at 456-57 (collecting U.S. cases).

92. See AREEDA, *supra* note 60, at 199; 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 735c, at 368, ¶ 740d1, at 431-32; AREEDA & TURNER, *supra* note 8, ¶ 712, at 155-56.

93. See Baumol, *supra* note 41, at 57 n.13.

94. See AREEDA, *supra* note 60, at 199-200; 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 740d1, at 432; Baumol, *supra* note 41, at 61-62.

95. Although acknowledging that the period of predation is the correct time period “theoretically,” 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 740d1, at 432, Areeda and Hovenkamp recommend adopting a “relatively arbitrary definition of middle-run variability” based on administrative concerns and a crude overall judgment that it is reasonable to deem “most

The basic logic is simple enough. Until the alleged predatory price lasts long enough to be exceeded by those costs that were *variable for that period*, an equally efficient entrant cannot have suffered any loss it could have avoided by exit, and thus cannot have had any incentive to exit. Alleged predatory prices that last only one month cannot cause an equally efficient rival to lose any money by not exiting unless those prices are lower than the very short-run costs the rival incurred by operating that month. In contrast, pricing that lasts for ten years will cause an equally efficient rival to lose money (relative to exit) if the price does not suffice to cover the fixed costs of producing anything next year (like overhead) or the future capital costs of rebuilding facilities that seemed like sunken costs in the short run but are variable over a time horizon of ten years. Thus, we need not pick one time period or cost measure in the abstract; the choice is dictated by the time period of the alleged predation.

One implication of this is that, for purposes of predatory pricing law, one should thus not distinguish between sunk, fixed, avoidable, and variable costs with general definitions about whether they are escapable in a limited period, or need to be incurred to produce any output or to produce anything beyond the first unit of output. Rather, the question of whether (and what) costs to consider should depend *solely* on whether they could be varied during the time period of the alleged predation. It may be that the costs that could be varied during this period include the costs of making the first unit, and thus include what Baumol calls avoidable costs. But it may be that the costs of making the first unit of output cannot be varied during the relevant period, and thus should be excluded. Indeed, it may be that the costs of making the first 100 units cannot be altered, in which case the costs that are variable during that period do not even include all costs that are traditionally lumped into average variable costs. At the other extreme, it may be that the period of alleged predation is so long that it includes what would be deemed sunk costs under a definition that considers whether they are inescapable over a limited period. I will thus call a cost “variable” if it could be altered during the period of alleged predatory pricing, and “fixed” if it could not be altered during that period.

What is the concern of those who favor using long-term costs, even when the predatory pricing period is short? One theory is that predatory pricing at the “rival’s variable costs” can induce their exit because “[t]he

costs” variable, *id.* ¶ 740d1, at 433-34; see also 3 AREEDA & TURNER, *supra* note 8, ¶ 715c, at 173-74. But it is not clear why the administrative concerns should be so great given that the time period of the alleged predatory prices is presumably known, nor why it should be reasonable to make an allocation that is clearly wrong for many time periods or firms. Further, any seeming administrative advantage from using a categorical definition seems eliminated by Areeda and Hovenkamp’s willingness to abandon average variable costs, or narrow or broaden their definition, when the categorical rule seems to lead to bad results. 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 740b3, at 429-30, ¶ 740d4-5, at 437-39, ¶ 741e, at 449-55.

rival, who also incurs fixed costs, exhausts its financial resources and leaves the market.”⁹⁶ But this is wrong. As long as the price exceeds the costs a rival could vary during the relevant time period, the rival would lose money from leaving the market. True, it might have to renegotiate loans or go into bankruptcy because it becomes unable to meet any loans that it incurred on the assumption it could cover its long-run sunk or fixed costs. But since the rival is worth more as a going concern (which follows from the assumption that prices exceed its variable costs), even then the business will be maintained in bankruptcy and the firm will stay in the market. This point is sometimes lost because of the popular image that firms somehow “vanish” in bankruptcy, but in fact bankruptcy reorganizations just change the owners of the business from shareholders to creditors, and the bankruptcy trustee as fiduciary for the new owners has the obligation to continue operating the firm if that creates profits for the new owners. (Note that the question would be different if below-cost pricing were inflicting actual losses, for then the firm would have to convince creditors to provide additional funding to keep the firm afloat.) Thus, predatory pricing at the costs that are variable to the rival may injure the rival’s shareholders or lenders, but cannot drive out an equally efficient rival. Accordingly, no rational predator would do it, especially since the predator would be inflicting the same injury on its own shareholders or lenders. There is also the question why the predator would have any better access to capital markets than the rival, but that is a general question for below-cost predatory pricing.⁹⁷ Here, the problem is that pricing above the costs that are variable to the rival cannot inflict any loss that drives out the rival at all, even if the predator *does* have better access to capital markets.

A related theory appears to be that an equally efficient rival would exit the market after even a short period of prices below long-run costs because the rival sees before it a future where prices will not allow it to stay in business. Thus, some think that when the predator prices at the rival’s variable costs, “a rational rival should leave at the first indication that the incumbent is even contemplating a predatory campaign, there being no point in sticking it out and squandering resources when exit is inevitable.”⁹⁸ But this too is wrong. Even if the rival were convinced the predator’s pricing will be permanent,⁹⁹ it would have no incentive to exit prematurely. Until the rival begins to have to make decisions on whether to keep

96. Ordovery, *supra* note 17, at 79-80 (summarizing the literature); see also Williamson, *supra* note 7, at 322 (accepting the deep pockets theory).

97. Ordovery, *supra* note 17, at 80.

98. *Id.* at 79.

99. There are reasons to doubt the predator’s ability to credibly commit to continue a scheme of pricing that imposes long-term losses on itself or even forgoes short-term profits, but those reasons are equally applicable to below-cost pricing. *Id.* (describing the objection that below-cost pricing is irrational because the predator cannot recoup its losses).

incurring fixed or capital costs, those costs will not be variable to it, and the rival will stay in the market because it is making a profit at the alleged predatory price relative to the costs it could vary during the period of such pricing.

Thus, a price that is below a rival's long-run costs will not cause it to exit the market as long as the price allows it to cover the costs that are actually variable during the period of alleged predation. To test the proposition, let us directly confront the example that has most bedeviled modern antitrust scholars. What do we do with software whose marginal or variable cost of production is near zero? The usual answer is that the "new economy" has to be treated differently because marginal or variable costs are so low.¹⁰⁰ But this creation of an ad hoc exception is hardly satisfactory. In the old economy, marginal or variable costs are also often below average or long-term costs. Indeed, the distinction between these cost measures only matters because they sometimes diverge. If this divergence presents a big problem when it is large, it must present at least a small problem when it is small. Our theory for how to deal with that divergence should be able to address the full range of possible magnitudes rather than having ad hoc exceptions, especially since those exceptions create ambiguity about just what the vague dividing line might be.

The better answer is, instead, that it all depends on how long the pricing lasts. If pricing at a near-zero level occurs for a short time, it cannot persuade any equally efficient software rivals to exit, since they also will have near-zero marginal costs and thus retain a profit from operating during that period. If instead such pricing lasts for years, then it could be predatory because it would not allow an equally efficient software rival to recoup the software development costs of updating that software to stay in the market. The latter costs become variable to the rival if the predatory pricing is lengthy, but not if it is brief. Paradox solved.

100. See, e.g., Bolton et al., *supra* note 15, at 2272-73. The authors base this on the assumption that "the short-run incremental cost of a program downloaded from the Internet is nil." *Id.* at 2272. This is probably overstated, since advertising and marketing effort affects the level of sales through downloads, and additional downloads require more billing effort and technical support. But the incremental costs do seem very low relative to the fixed or sunk costs of making the software, providing a sound basis for their conclusion. Likewise, Areeda and Hovenkamp are sympathetic to cases that make an exception to average variable cost rules in regulated or high-technology markets (a) with "an unusually high ratio of fixed to variable costs," 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 741e2, at 452, or (b) where the industry is "expanding," *id.* ¶ 741e2, at 454. See also *id.* ¶ 741e2, at 452-55 & n.48 (agreeing that long-run incremental costs are relevant in such cases, but ultimately deciding that the test cannot be implemented). Factor (a) is irrelevant for reasons noted in the text. Factor (b) is relevant to the extent it means the capital costs are in fact variable during the period of the alleged predatory pricing. *Cf. infra* text accompanying notes 110-113 (noting that when an industry is contracting, capital costs may not recur). Thus, rather than creating a special exception, it is more straightforward to see this as one application of the general principle.

The fact that marginal or variable costs are *uniformly* low in an industry thus raises no difficulty if one is careful to consider all (and only) costs that are variable during the period of alleged predation. Different problems might be raised, however, if the equally efficient rival has a different ratio of fixed to variable costs than the alleged predator, or if the alleged predatory price is timed after the predator has incurred a fixed or sunk cost that the rival must decide whether to incur in the future. I address those issues next.

C. *Use Variable Costs of the Alleged Predatory Increase in Output That Displaces the Rival, Not of Producing the Predator's Entire Output*

Another common concern is that equally efficient firms might have different ratios of fixed and variable costs. For example, Williamson observes that more capital-intensive firms can have lower variable costs even when they are less efficient than more labor-intensive firms.¹⁰¹ He thus advocates using average total costs as a better means of sorting out the efficiencies of firms.

This is a reasonable concern with using the average variable costs of making the predator's *entire* output because that measure is by definition lower than average total costs.¹⁰² Thus, if allowed to price at this measure of average variable costs, even a firm exceeding its optimal capacity could price at a level that is lower than its marginal or average total costs, and thus lower than the costs of an equally efficient firm at providing that incremental output. An average variable cost test can thus offer inadequate protection to an equally efficient rival *if* it is based on an average of the costs that are variable for the predator's entire output.

But this does not mean that one must abandon use of a variable cost test altogether. It simply means one must be more precise in defining the relevant output whose costs can be varied. Since our purpose is to determine what cost measure would prevent a firm from excluding an equally efficient rival, the relevant costs that are variable are not the costs of producing the predator's entire output. They are rather the variable costs of the alleged predatory *increase* in output that displaces the *rival's*

101. Williamson, *supra* note 7, at 321-22. In fact, a capital-intensive firm may not always have lower variable costs. Rather, variable costs depend, in part, on how much the firm's capital assets depreciate with increased use. 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 735c2, at 370-71.

102. By definition, average total costs exceed average variable costs because average total costs are the sum of average fixed and variable costs. See *supra* note 19.

output.¹⁰³ This is because the concern is rival exit (or nonentry), and thus the question is which firm is more efficient at producing the rival's output.

In other words, the relevant concern is that the predatory price allows the predator to expand output so that the additional output displaces the rival's equally efficient output and thus precludes it from the market. The relevant cost measure is thus whatever, over the period of alleged predatory pricing, is the cost of producing that higher output minus what the cost would have been of producing the lower output. The higher output will be the alleged predator's output at the predatory price, and the lower output will be that output minus the rival output that was allegedly displaced by predation. (Where the rival was already in the market, the lower output will normally equal the predator's output before the alleged predatory pricing began.) This measure of variable costs, in effect, is the sum of the marginal costs for the predatory increase in output, but can be measured more simply by comparing the costs at the higher output to the costs at the lower output, rather than by trying to calculate the marginal costs of producing the last item at each output level. Assuming marginal costs are increasing, this total variable cost figure (when divided by the increased output to give a per-unit figure) will give an average variable cost that is lower than the marginal cost of producing the *last* item that the alleged predator makes, but will still protect a rival that is equally efficient at making the relevant increment of output.¹⁰⁴

103. See Baumol, *supra* note 41, at 64-65.

104. Any price below the predator's marginal cost of producing the last unit of output would still be inefficient for at least some of its increased output, but would not drive out an equally efficient rival. If the predator is expanding beyond its optimal capacity, a price above its variable cost of increasing its capacity must be above its minimum average total costs, and thus cannot displace production by an equally efficient rival. Instead, the monopolist simply will inefficiently confer a boon of extra output at below-cost prices to buyers. Since the rival cannot be driven out by such pricing, the monopolist will not recoup any losses, and thus such inefficient pricing should be self-detering.

The analysis grows a bit more complicated if the putative predator is operating *below* its optimal capacity. Suppose that before the predatory behavior has begun, the predator is below its optimal capacity by the quantity of output (Q) that could replace the victim's output. It then increases output to its optimal level, but instead of charging its marginal cost, it charges a price equal to its average variable cost of that increased output, which is somewhere above its marginal cost at the original output and below the marginal cost at its final (optimal) output. Some of this increased production has a marginal cost that exceeds its price, and one might thus worry that this is a predatory increase in output that could replace the output of an equally efficient rival, whose average variable cost of production is greater than this price but lower than the predator's total cost or the marginal costs of its last unit of production. But the rest of the predator's increased production is efficient (and cheaper than the rival's), and the overall increase in predator output cannot displace the rival's production of the *same* output unless the rival has higher variable costs of producing it, in which case the rival is not equally efficient at producing that increment of output. The solution to this apparent anomaly is that the proper predation claim is not that the entire increase in output was predatory. Rather, the predatory portion of the output increase was where the output's marginal cost began to exceed the price. For that properly defined increment of output, the variable cost of providing it will be above price.

Accordingly, if the capital-intensive firm has increased output to displace its rival's output, we should look only to the higher variable costs of the allegedly predatory increase in output, not to the lower variable costs of producing the predator's entire output. Prices at or above those higher average variable costs cannot drive out a rival that is equally efficient at making that increment of output. If the capital-intensive firm's variable costs of increasing its output enough to displace the rival are lower than the rival's own variable costs of producing that output, then the rival is in fact not equally efficient at making its output. Rather, the rival output can more efficiently be supplied by an increase in the capital-intensive firm's output, even though it may be exceeding its optimal capacity.

This approach resolves a conundrum created by the approach of Areeda, Turner, and Hovenkamp. Although their cost measure means a predator should shut down when its price is lower than the average variable cost of its entire output, they recognize that this creates an anomaly when there is so much excess capacity that this legal rule would require every firm in an industry to shut down, and thus they create another exception to their own rule.¹⁰⁵ But under the equally efficient entrant benchmark, the question is not whether the predator is profiting by producing its output. Rather, the question is whether it could profitably displace the *rival's* output. A firm pricing at marginal costs that are below its overall average variable costs necessarily lowers those average variable costs by expanding output. Thus, the fact that its prices are below its overall average variable costs does *not* mean they would be below the additional variable costs it would incur by adding output equal to what the rival used to produce. In such a case, the declining demand that created the excess capacity simply means that the minimum efficient scale can sustain fewer firms than before. Firms that can produce the rival's output more cheaply than the rival should be able to price down to the variable costs of increasing their output, even

To take a concrete example, suppose the predator's $MC = Q$, and it begins at an output of 90 and a marginal cost of \$90 and subsequently increases this output by 10 to reach its optimal capacity of 100, at which its marginal cost is \$100. But instead of charging \$100, the predator charges \$95/unit, which fully covers the variable cost of this output increase of 10. To test the proposition that it will not matter if a higher proportion of the labor-intensive rival's costs are variable, take the case where all of the rival's costs are variable and equal \$96/unit, reflecting constant labor costs. One might be concerned that by pricing at \$95 and increasing output by 10, the predator would be able to drive this rival out of business even though, at \$96/unit, its costs are lower than the capital-intensive firm's minimum average total costs and marginal costs. But if the rival has constant variable costs of \$96/unit, then the predator, by pricing at marginal cost, would not have produced 90 units at \$90, but rather 96 units at \$96/unit. Thus, the rival would only have had 4 units of output to replace, and the variable cost of replacing those last 4 units would be \$98. Since those are the relevant variable costs, the predator could not price below \$98 for those 4 units, and thus could not drive out the rival under a properly defined cost test if it were correctly alleged that the predatory increase in output consisted of the final 4 units. The first 6 units were replaced by more efficient production and thus should not be part of the predation claim at all.

105. 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 740b3, at 429-30.

when these fall short of their overall average variable costs. This output expansion and displacement of rivals will allow the surviving firms to reestablish their optimal capacity and rebalance supply and demand.

D. *If Short-Term Pricing Can Deter Long-Term Investments,
Then Use Magnitude of Predator Costs for the Sorts of Costs
Variable to the Victim, but Look to the Future
To Measure Cost Magnitudes*

Another concern, which the literature has neglected, is that the predator might time its alleged predatory pricing to begin after the predator has incurred a sunk cost, but right before its rival has to decide whether or not to do the same. Suppose, for example, the predator has just renewed a ten-year lease on its factory, but knows that its equally efficient rival has an upcoming decision about whether to renew its own factory lease. The predator then cuts prices to a level that does not suffice to cover the sum of operating and lease costs. The price exceeds the predator's variable costs since its lease costs are sunk.¹⁰⁶ But the price does not exceed the variable costs of the equally efficient rival because the rival can avoid committing to the lease and thus its lease payments are variable at this time. If that market price persists, the rival will lose money by renewing the lease and thus has incentives to exit the market rather than renew. The same holds if a firm lowers software prices to near zero after it has come out with the latest software update, but before its rival has followed suit by investing in its own software development.

Now, there are good reasons not to treat this concern as serious because short-term pricing probably cannot deter long-term investments. But let me defer those reasons until the next Section. Here, the point I wish to focus on is that, even if this concern were serious, it still would not justify a general rule of always employing long-term or average total costs. Rather, the solution lies in defining more carefully just which variable costs one examines.

Because the goal is to make sure our cost measure is not protecting less efficient firms, antitrust examination normally focuses on the *predator's* variable costs. It would be more precise to say that this benchmark requires that *the magnitude* of any variable costs come from the predator's cost data. The *sorts* of costs that should be deemed variable would, if this concern

106. The lease cost is sunk for that ten-year period even though the rents will be paid in the future because the obligation to pay them will exist whether or not the firm stays in production. I simplify the situation here for exposition—in fact, there is probably some possibility of a sublease, just as there is some alternative use for just about any sunk investment. The actual sunk lease cost is thus, more precisely, the difference between what the company must pay on the lease and what it might get with a lower-valued sublease or substitute use.

were serious, instead turn on whether those sorts of costs were variable *to the rival* during the period of alleged predation. The reason is that the purpose of our test is not to determine whether the price is profitable to the alleged predator in the short run, but rather to determine whether it could drive out an equally efficient rival.

The following hypothetical illustrates the proper approach if this concern were serious. Suppose that each of two firms has one plant that costs \$10,000 a year to lease and makes 1000 gizmos annually. Take two factual scenarios. In the first, each firm also has operating costs of \$10 per gizmo and is thus equally efficient. The alleged predator then leases two plants and begins charging \$15 per gizmo. If the rival has already rented its plant too, the costs of leasing a plant are not variable to the rival. The correct measure of variable costs is thus \$10 per unit and the price is not predatory because it cannot cause the rival to exit. If the rival has not rented its plant yet, its variable costs include not only its operating cost but also the cost of leasing a plant. The correct measure of variable cost is thus no lower than \$20 per unit, and the same \$15 price is predatory because it inflicts a loss on this rival that might cause it to exit. This difference in results is not anomalous because the \$15 price can drive out the rival that has not incurred a lease obligation but cannot drive out the one that has because, although equally efficient, each rival will compare the \$15 price to the differing costs that it can vary.

Now consider a second scenario, where the rival has operating costs of \$15 per gizmo and is thus less efficient. If the rival has not rented a plant, we should look to both lease and operating costs because they are the *sorts* of costs that are variable to the rival. However, \$25 is not the right measure of those costs because the *magnitude* of those sorts of costs must be determined by looking at the predator. Although any price below \$25 could inflict losses that might drive out this rival, that is true only because it is less efficient. The right cost measure is \$20 per gizmo, reflecting the magnitude of the alleged predator's costs for the sorts of costs that are variable to the rival.

A related concern is that, even if a properly defined measure of variable costs can prevent an incumbent from *driving out* an equally efficient rival, it may not prevent an incumbent from *detering entry* by equally efficient firms. An entrant, this concern stresses, will not enter unless it expects prices to cover its sunk costs of entry. Thus, if an equally efficient entrant anticipates incumbent price levels that cover variable costs but do *not* cover sunken entry costs, it will not enter.¹⁰⁷ But if this concern is a serious one, a question I will take up below, it really is no different than the last case.

107. See Easley et al., *supra* note 78, at 447-54 (offering a model under which an incumbent could deter entry with prices above average variable costs).

Deterring entry is just a special case of deterring sunken investments. The correct time period in this case starts before entry because that is the period during which the rival decision is influenced. The sorts of costs that are variable during this period to the potential entrant include the (not-yet-sunk) capital costs of creating new facilities. Thus, the correct cost measure should include the incumbent's (amortized) long-term capital costs of replacing its facilities with new ones. If the incumbent's future costs of plant replacement are lower than the entrant's cost of building its plant, then the incumbent should be able to manifest that greater efficiency in lower long-term pricing even though it excludes less efficient entrants.

Even when the rival's cost variability during the relevant period indicates including the above sorts of capital costs but measuring them by their magnitude to the incumbent, this does not mean we should look to the incumbent's average total costs. The problem with most measures of average total costs is that they look *backward* at the sum of variable and fixed costs the firm has already incurred. But what matters (if this concern is serious) is the magnitude of the *future* costs the incumbent will incur if its alleged predatory pricing persists. If the market is in a steady state, then basing average total costs on past data is a good proxy for future long-run costs. But the proxy might be poor if the market is changing. For example, if the industry is declining, then such measures of average total costs are a poor proxy because firms should be contracting or exiting, and thus their past sunk or fixed capital costs will not recur. Combining this future orientation with the other analysis above helps address a nagging debate when the alleged predator's marginal costs are lower than average total costs.

Defenders of marginal or variable cost measures have tended to stress that, if short-run marginal costs are below average total costs, then by definition expanding output should lower average cost, which must mean the alleged predator is below its optimal (least average cost) output and has excess capacity.¹⁰⁸ They thus conclude it will be cheaper to use that excess capacity than to build new, more expensive capacity. Objectors have tended to stress that marginal or variable cost measures of predatory pricing give monopolists inefficient incentives to build the excess capacity that is

108. See, e.g., AREEDA, *supra* note 60, at 195-96; 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 724c1, at 287, ¶ 739b, at 414-15, ¶ 741d1, at 446-47; 3 AREEDA & TURNER, *supra* note 8, ¶ 715a, at 164-68 & n.7. Actually, there is technically one exception: It might be the case that the marginal cost of the final item produced is below average total cost, but that the cost of adding one more unit of output would exceed average total cost. For example, in the lease hypotheticals noted above, the strict plant output limit of 1000 meant that going from 2000 to 2001 units has a marginal cost of \$1010. Thus the \$10 marginal or variable cost of making 2000 units is below the average total cost of \$20 even though the predator is not below optimal plant size and does not have excess capacity. But if output limits are less strict, marginal costs will rise less sharply and this exception will not arise.

necessary to justify the future pricing below average total costs that makes entry unprofitable.¹⁰⁹ Applying the approach outlined above can provide a more systematic resolution to the problems raised by the four typical sources of excess capacity.

- (1) We might have a declining industry. Here we would not want to require prices that cover capital costs because that would encourage investment and entry at a time when market economics dictate exit. Advocates of total cost measures, like Posner and Williamson, have tended to respond by creating a declining-industry exception to their favored cost measure.¹¹⁰ But a more satisfactory answer is again to be more precise about which costs we are measuring and when we are measuring them, rather than to use overinclusive cost measures or make equally overinclusive exceptions. To the extent that plant-replacement costs will *not* recur in the future because firms are contracting or exiting, then the future magnitude of those predator costs will be zero.¹¹¹ The incumbent's future capital costs will thus be far below the past capital costs reflected in backward-looking measures of average total

109. See, e.g., Posner, *supra* note 89, at 942; Scherer, *supra* note 7, at 871 n.12; A. Michael Spence, *Entry, Capacity, Investment and Oligopolistic Pricing*, 8 BELL J. ECON. 534 (1977).

110. See POSNER, *supra* note 56, at 189; Williamson, *supra* note 7, at 322-23. They are not the only ones who create an ad hoc exception in this circumstance. Areeda and Hovenkamp also create an exception to their rule banning prices below average variable costs when this results from industry-wide excess capacity. See *supra* text accompanying note 105 (explaining how that issue can instead be addressed by defining the output whose costs are in question). They also create a similar exception when a defendant builds a plant that turns out to be so costly compared to demand that prices do not cover a standard measure of average variable costs that includes use depreciation. 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 740d5, at 437-39. Rather than creating a special exception for this case, it is clearer to see it as just one result of the general rule that, when the sunk costs of building the plant will not be incurred again, the future magnitude of any such costs (whether manifested in use depreciation or otherwise) is zero. Areeda and Hovenkamp recognize a theoretical exception in the other direction when prices exceed average variable costs but the industry is growing and fixed costs greatly exceed variable costs. See *supra* note 100 (explaining how that issue can instead be resolved by assessing whether the costs are variable during the time period of the alleged predation). They also create an exception when prices exceed average variable costs, but fixed costs were incurred just to drive out the rival. See 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 741e1, at 449-50. Again, it is clearer to see this instead as just one application of the general rule that the relevant variable costs are those that the firm can vary during the period of predatory pricing in order to create the additional output that replaces rival output.

111. One might think that predator costs can never be zero because the predator will have to replace its plant at some point to stay in the market. But there are two possibilities. First, the predator might have multiple plants. Then, what matters is the long-run cost of operating the marginal (least efficient) plant(s) that can replace the victim's output. In the face of declining market demand that produces prices that do not suffice to cover the capital costs of plant building, the predator will close the marginal plant(s) rather than rebuild it (them), and thus it (they) will have zero future capital costs. Second, the predator might have only one plant. Here, if demand has declined to the extent that the predator can efficiently supply the entire market with this one plant, then the decline has made the predator into a natural monopoly and the analysis that follows in the text for natural monopolies would apply.

costs.¹¹² By instead combining the magnitude of future predator costs with the sorts of costs that are variable for the potential entrant during the period of predatory pricing, we can produce the right result without having to patch up a hole in the cost measure theory with an ad hoc exception.¹¹³

- (2) We might have a temporary cyclical decline in demand that creates temporary excess capacity. But since one cannot know whether the decline is temporary until it ends, during any demand dip the magnitude of the alleged predator's plant-replacement costs should be zero, as above. Pricing at that level will defer entry, which is the right result since during that time the entrant will be less efficient than a firm that need not incur capital costs. But entry will not permanently be deterred if the decline is indeed temporary.¹¹⁴ Nor will the predator be able to drive out any equally efficient existing firm that also has excess capacity with any price that exceeds their (equal) variable costs during that temporary period.¹¹⁵
- (3) We might have economies of scale that make it cheaper for the largest firm to provide additional output. But once again we do not have to create an ad hoc exception. Even if plant-replacement costs are the sorts of costs that should be considered variable, their magnitude is determined by the incumbent's costs. Given economies of scale, the incumbent monopolist would incur smaller such costs in future production of the output that the entrant proposes to add than the entrant would. Pricing at those future incumbent costs will deter entry, but the entrant is not equally efficient given the relevant economies of scale.¹¹⁶
- (4) The incumbent monopolist might be retaining inefficient excess capacity on hand in order to be able to respond to entry. Defenders of marginal or variable cost measures acknowledge

112. It is surprising that Posner did not make this point himself, since he so insightfully pointed out the past-future divergence between average total costs and future marginal costs. See POSNER, *supra* note 56, at 190.

113. Alternatively, one could say that what matters are "anticipated" average total costs, and that neither they nor "long-run marginal costs" or "long-run incremental costs" should include capital costs that will not recur.

114. Suppose we instead assume that it is crystal clear the demand decline is temporary, but that entry cannot be deferred. That is unlikely, but if so, any entrant would know to discount the temporary decline in demand and enter now, recovering entry costs when demand returns if entry is efficient.

115. Areeda and Turner suppose that it might, but do so based on the argument about exhausting rival financial resources that was rebutted above. Compare *supra* Section II.B, with 3 AREEDA & TURNER, *supra* note 8, ¶ 715a, at 165-66.

116. For more extensive analysis of when the entrant will be equally efficient given declining costs, see *infra* Subsection IV.C.3.

the difficulty, and either advocate an exception or reject one as inadministrable and accept this as a downside of their rule.¹¹⁷

But a more satisfactory approach would recognize that in such a case the magnitude of future incumbent capital costs will include the replacement costs of maintaining that inefficient excess capacity even if the market is in steady state. And that sort of cost must be included because that is the sort of cost that is variable to the entrant. A monopolist required to price at the requisite cost level thus will not be able to keep out an equally efficient entrant even if the entrant believes the pricing will persist indefinitely, and will thus have no incentive to create excess capacity in order to make such an attempt.

E. *If (as Likely) Short-Term Pricing Cannot Deter Long-Term Investments, Then Just Use Those Predator Costs Varied by Its Alleged Predatory Increase in Output*

The preceding Section assumed the concern that short-term predatory pricing might deter a long-term investment was serious, a proposition that is doubtful for reasons I now take up. The main problem is the following: The claim that pricing (or threats to price) above the alleged predator's variable costs might deter investment or entry by equally efficient firms depends on a crucial supposition. That supposition is that, in making its long-term investment or entry decision, the rival will believe that such pricing will *persist* in the long run, or that any threat to impose such pricing after entry will both be carried out and persist in the long run. This supposition is what allows an alleged predator with a *short-term* pricing strategy (or mere threat to begin such pricing) to influence rival investment or entry decisions that are made based on *long-term* expectations. But this supposition is dubious because any equally efficient rival will realize that, if it incurs the sunk cost in question, it will no longer be rational for the alleged predator to persist in pricing that covers variable costs but not sunken capital costs, let alone to carry out a threat to begin such pricing. The reason is that, once the sunk cost is incurred, such pricing cannot give the equally efficient rival any incentive to leave the market. Since the alleged predator could make more money with pricing that covers these long-run costs, and cannot drive out the rival with lower pricing, it would be irrational for the predator to persist in such low pricing. The prospect of such irrational pricing thus would not induce exit by the equally efficient rival, which would instead assume that any unremunerative pricing would not continue.

117. AREEDA, *supra* note 60, at 198-99 (rejecting the exception); 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 741d2, at 447-49 (suggesting both); Joskow & Klevorick, *supra* note 71, at 253-54 (recognizing one).

This might seem indistinguishable from the claim that an equally efficient rival cannot be deterred or driven out by pricing *below* the predator's variable costs because it would be irrational for the predator to persist in such money-losing pricing.¹¹⁸ But actually the issue here is different. Pricing below variable costs inflicts actual losses on an equally efficient rival that *would* induce the rival to exit if it believes the predator will persist, which itself might make it rational for the predator to persist in order to drive the rival out. Here, once the rival incurs the sunk cost, pricing above variable costs cannot inflict any future loss on the rival and thus cannot give it any incentive to leave the market. A single-market monopolist will thus have no reason to persist in such a pricing strategy.¹¹⁹ It will instead raise prices to whatever level maximizes profits given that the rival cannot be eliminated.

This means the incumbent monopolist who cannot price below variable costs will have strong incentives to price even *higher* than long-run costs in response to an equally efficient entrant. Because the entrant has committed the sunk costs, the monopolist cannot drive the entrant out with any low price that is above their equally efficient variable costs. The addition of the entrant has thus converted the former monopoly to an unavoidable duopoly. Thus, as soon as it realizes the entrant is equally efficient, the incumbent monopolist will endeavor to accommodate entry by pricing at supracompetitive duopoly levels rather than dropping prices to less profitable above-cost levels in a fruitless attempt to drive out the entrant.

This has the interesting implication that, if not permitted to price below its variable costs, the incumbent monopolists *themselves* will want to sort out the equally efficient entrants from the less efficient ones as accurately as possible, and will only attempt to drive out the latter with price cuts. If the incumbent monopolist does react to entry with a price cut that is below its long-run costs but above the costs of replacing the victim's output that are variable to it during the period of alleged predatory pricing, then it must believe the entrant is not equally efficient. Further, if that short-run price cut fails to drive out the entrant, then the incumbent monopolist will realize in the long run that its belief about the entrant's relative inefficiency was mistaken, and endeavor to raise prices back to a supracompetitive duopoly level. Given this long-run prospect, short-term price cuts to levels that are above the incumbent's variable costs should not deter investment or entry by equally efficient firms.¹²⁰

118. See *supra* text accompanying note 17.

119. The perhaps counterintuitive implication is that, where the incumbent and entrant are equally efficient, pricing just barely above variable costs, but below long-run costs, is actually a *less* rational predation strategy than pricing below variable costs. Cf. Edlin, *supra* note 6, at 961-63 (assuming that above-cost predatory pricing must be more rational than below-cost pricing).

120. Where the entrant is *more* efficient than the incumbent, it may instead prefer to persist in prices that are above its variable costs, but below the incumbent's variable costs, in order to drive

True, one could try to extend some of the countertheories used to justify bans on below-cost predatory pricing to this case where predator prices do not cover its long-term costs. One theory is that the predator's bluff to persist in such pricing may never be called because the short-term predatory pricing deters investment or entry by misleading the rival (or capital markets) into thinking the predator's efficiency is greater (or market demand lower) than it actually is.¹²¹ But this would not apply to a mere threat to lower prices in response to investment or entry. Actual lower prices would be necessary to create the misleading impression. This theory thus has little application to the topic of price cuts reactive to new entry.¹²² Nor would past reactive above-cost price cuts have much future reputational effect in the same market. Unlike a price cut below variable costs, a price cut above those costs cannot drive out an equally efficient entrant who has incurred the sunk costs of entry, and thus will eventually be abandoned and fail to create the impression that the incumbent is more efficient than entrants. Further, in the present context, the assumption that other firms and capital markets can be fooled in the long run seems dubious. Such pricing creates a market opportunity for any capital investors savvy enough to realize when current prices are an unreliable indication of future prices, especially since the actual future prospects are that the equally efficient entrant will get a share of supracompetitive profits.¹²³ More important, even if actual short-term pricing that did not cover long-term costs does fool rivals (and their providers of capital), their investment or entry will only be *deferred*. To continue deterring it, the predator will have to maintain such pricing for the long term. If it does so, then such pricing below long-term total costs will become predatory because the relevant capital costs will have become variable for the predator too during the long

the incumbent out of the market and become the new monopolist itself. But that prospect will hardly deter entry.

121. Ordoover, *supra* note 17, at 80-81 (synthesizing the recent literature); see also Bolton et al., *supra* note 15, at 2247-49, 2285-330 (same but in greater depth).

122. On the other hand, if the entry is announced, but not yet completed because some important capital investment remains to be made, see *infra* Section V.A (discussing the difficulties of defining the moment of entry), then price cutting at that stage might deter the investment necessary to complete entry if it misleads the entrant. But it is hard to believe entrants will be that misled by pricing that is plainly reactive to their entry plans.

123. Any assumption about uncertainty must also be applied evenhandedly. The predator will also be uncertain about entrant efficiency and future consumer demand. If the rival is less efficient, pricing below total costs would sacrifice profits for no good reason since total cost pricing would deter investment or entry anyway. If the rival is more efficient, then pricing at incumbent variable costs may not deter investment or entry even if the entrant mistakes that price for an indication of the incumbent's total costs. If the rival is equally efficient, the predator will still be uncertain whether the rival (and its providers of capital) will interpret its pricing as indicating total rather than variable costs. Uncertainty about what sort of rival or potential rival it faces, and about how any price signal will be interpreted, thus seems sufficient to deter the predator from pricing at levels that sacrifice profits in all cases in order to send a signal to a mere subset of rivals, especially since that level of pricing cannot in fact inflict post-investment or post-entry losses on equally efficient rivals.

period of alleged predatory pricing. Thus, even if one believes capital markets are easily misled, that is no reason to deviate from using the costs that are variable to the predator during the relevant period.

Alternatively, one might conclude it is rational for the predator to persist in pricing that does not cover long-term costs in one market if it is a monopolist in *many* markets and wants to signal firms in other markets that they will lose money if they enter those other markets (or incur the periodic sunk costs necessary to stay in them).¹²⁴ For example, suppose that after an equally efficient firm enters one market, the incumbent responds with a price that allows the entrant to recover the costs that are now variable to it, but that does not allow recoupment of its sunk costs of entry. Such a price cannot drive out the entrant for reasons described above and would thus be irrational if only the first market were considered. But suppose the incumbent does not set its price to drive out the first entrant. Instead, it sets that price to deter other equally efficient firms—who have not yet incurred entry costs—from entering the other markets. If the other potential entrants believe the incumbent will respond with the same pricing in those other markets, they will be deterred from entering (even though equally efficient) because they cannot recoup their entry costs. A similar strategy might be employed to deter the sunken investments necessary for existing rivals to stay in multiple markets.

Or so goes the theory. But there are manifold problems with this multimarket theory of predation through prices above variable costs. The first is obvious. Often the alleged predator is not a monopolist in multiple markets, making this theory utterly inapplicable. Second, it will rarely be the case that in *all* the predator's markets, the predator has made sunk investments about which rivals or potential entrants are just about to decide. Such a strategy thus cannot help induce exit or deter entry in any markets where the rivals have already incurred the relevant sunk cost. Third, this pricing strategy cannot deter investment or entry by any rival that simply invests in or enters all the remaining markets simultaneously, since then the pricing cannot send a signal to any remaining market.¹²⁵ Capital markets should be willing to provide the financing to increase the scale of entry because getting (or retaining) a slice of supracompetitive profits in these markets will be highly profitable. And if no single rival can invest or enter in all markets, multiple rivals or entrants can always organize a group of firms to make simultaneous investments or entry.¹²⁶

124. See Easley et al., *supra* note 78, at 447-54, 457 (offering a multimarket model under which an incumbent could deter entry of equally efficient entrants in subsequent markets with prices above variable costs in the first market); Ordober, *supra* note 17, at 80 (reviewing the literature).

125. Easterbrook, *supra* note 17, at 286-87.

126. *Id.* at 288. Since, by definition, the firms would be in separate markets and unable to enter them all, they would not be horizontal competitors subjecting their agreement to judicial

Fourth, even if rivals cannot act simultaneously in multiple markets, such a pricing strategy cannot deter investment or entry by a rival in the *last* of the markets where rivals have not yet incurred the relevant sunk costs. The reason is that carrying out and persisting in such pricing will be irrational because it can neither drive out the last entrant nor send a signal in any future market. Since the threat is not credible, investment and entry by that last rival will not be deterred. Further, the rival in the next-to-last market would likewise not be deterred because the rival would realize such predator pricing would be irrational since it could not deter investment or entry in the only remaining market. And so on, until by backward induction one reaches the conclusion that the threat of initiating or continuing such pricing could not deter investment or entry in any of the prior markets.¹²⁷

In the case of below-cost predatory pricing, some have argued that backward induction fails because rival information is imperfect about whether incumbents can profit from below-cost predation against an equally efficient entrant.¹²⁸ But here that uncertainty is inapplicable since pricing above variable costs can never profitably drive out an equally efficient entrant. Others argue there is no clear end point at which a rival will know it is in the last market.¹²⁹ But applying this assumption evenhandedly implies equal ambiguity about who is in the *first* market that begins this supposed signal-sending game. If a predator is in ten ongoing markets and deprives a rival in only one market of the ability to recoup total costs, rivals in other markets seem more likely to draw inferences from the behavior in the nine markets than in the one outlier. The predator may thus need to carry out such a scheme in most markets to send a message to those that remain, which makes the scheme less rational (since profits will be sacrificed in a majority of markets where driving out the equally efficient rivals is impossible) and makes it clear to the remaining rivals that they are the last ones (which strengthens backward-induction problems).

More important, for the signal sent from any single market to be convincing, the predator will have to persist in the low price long enough to actually deprive its rival of a profitable long-term return on its investment

hostility. In any event, since an agreement to make simultaneous investments or entry need not involve any agreement on price, nor any agreement to refrain from entering or investing in each other's territories, it does not seem to involve any per se violation. And under the rule of reason, an agreement to add output without more would be procompetitive.

127. Reinhard Selten reaches the same conclusion for a threat of unprofitable below-cost predatory pricing. Reinhard Selten, *The Chain Store Paradox*, 9 THEORY & DECISION 127 *passim* (1978). If that conclusion holds there, a fortiori it will be true when variable cost pricing in the last market cannot inflict any loss that will induce the rival in the last market to exit after it incurs the sunk costs of investment or entry.

128. See Ordover, *supra* note 17, at 80 (reviewing the literature). Even in these models, there will be an equilibrium where a below-cost pricing strategy is only sometimes credible enough to deter investment or entry.

129. See *id.* (reviewing the literature).

or entry. If the predator just offers a price at variable costs for a short time, then it will not send the necessary signal that the predator is willing to persist in pricing below its total costs long enough to deprive an equally efficient rival of any ability to recoup sunk costs even when the rival cannot be driven out of the market. But the need to persist in such a scheme over the long haul to create an object lesson for other markets creates two serious problems. The first is that, by the time the incumbent has persisted long enough to create the signal in the first market, rivals will likely have made sunken capital investments in the other markets (which presumably share the same rate of capital replacement). Second, and more important, the predator's own capital costs will become variable during such a lengthy period of predatory pricing, and thus such pricing would be illegal under a variable costs test. Thus, *any multimarket predation plan by a monopolist must begin with conduct that would be an antitrust violation under variable cost measures* in at least the first market, and probably in most markets, in order to send the necessary signal. The imposition of treble damages in those markets should suffice to deter such a scheme.

If this analysis is correct, then it greatly simplifies the cost inquiry. Courts need not determine marginal costs or make complex judgment calls about which costs should be considered variable and which fixed and when to use one cost measure over another. Nor need courts determine the magnitude of the predator's costs for the sorts of costs that are variable to the victim during the period of alleged predatory pricing, which may entail capital costs and thus require projections about what sorts of capital costs the incumbent will incur in the future and what their amortized magnitude would be.¹³⁰ Instead, the relevant incremental costs are simply the difference between the actual total costs the incumbent incurred during the period of alleged predation and the total costs it would have incurred without the alleged predatory increase in output. Unless there has been some exogenous increase in input costs, this can often be determined by simply comparing total costs before and after the alleged predatory behavior. Dividing this by the alleged predatory increase in output converts this into a per-unit incremental cost, which then simply can be compared to the per-unit price the predator charged during the alleged period of predation.

F. *Conclusion on the Proper Cost Measure*

In short, it seems implausible that a predator could deter long-term investments or entry by any equally efficient firm by use of short-term threats or pricing strategies that exceed short-term costs. And when the

130. See *supra* Section II.D.

predator pursues a long-term pricing strategy, the difference between variable and total costs disappears because all costs are variable over the long term. My own conclusion is thus that allowing alleged predators to price at their own variable costs will not deter or drive out equally efficient rivals as long as we are careful to consider all costs of the allegedly predatory increase in output that replaces the rival's output that are variable *to the predator* during the period of alleged predation. That is, we need only consider those predator costs that are varied by the allegedly predatory increase in output. Accordingly, prices above this properly defined variable cost level should not be deemed predatory.

If the logic behind that conclusion were rejected, it would still be the case that a predator could not deter or drive out an equally efficient rival if its prices covered a cost measure reflecting the *magnitude* of predator costs for the *sorts* of costs in replacing the rival's output that are variable *to the rival* during the period of entry or investment decisions influenced by the short-term existence or threat of such pricing. Thus, even on this somewhat less sanguine view, prices above this somewhat higher cost level should never be deemed predatory even if below long-run total costs. A fortiori, prices above long-run total costs should not be predatory on any view, since everyone acknowledges they cannot exclude equally efficient rivals.

I should emphasize that the conclusions in the remainder of this Article hold regardless of whether I am correct about which cost measure suffices to assure that prices at cost cannot deter or drive out an equally efficient rival. Even if the reader disagrees with my above analysis about what measure of costs satisfies this standard, the analysis below would support rejecting a restriction on any price that is not below whichever cost measure the reader believes does satisfy this standard. That is, for purposes of establishing my general thesis, one can below substitute for the word "costs" whichever measure of costs the reader believes suffices to prevent an incumbent pricing at cost from deterring or driving out equally efficient entrants. While the lowest possible cost measure that satisfies this test may remain a matter of debate, there is consensus in the literature that a price at or above long-run incremental cost cannot drive out an equally efficient rival.¹³¹ Since the proposals to restrict above-cost price cuts would all ban

131. The European Advocate General expressly agreed with the general standard that predatory pricing law should favor "more efficient firms" and protect only firms that were "equally [efficient as] or more efficient" than the dominant firm. Joined Cases C-395/96 P & C-396/96 P, *Compagnie Maritime Belge Transps. SA v. Commission*, 2000 E.C.R. I-1365 ¶¶ 117, 132 (Opinion of Advocate Gen.). But he was of the mistaken view that selective above-cost price cuts could somehow drive out an equally efficient firm because of "its lesser financial capacity." *Id.* ¶¶ 122, 132, 138. In fact, this is impossible if one defines costs correctly, and certainly if one defines them to include all long-run marginal costs. Perhaps the Advocate General had in mind the intuition, shared by many theories, that a firm might be equally efficient in the long run, but not in the short run, and thus might need financing to overcome its initial inefficiency. I address that possibility below. See *infra* Section IV.C.

some prices above long-run incremental costs, they can be described as banning above-cost predatory pricing no matter which cost measure one uses.

III. REACTIVE PRICE CUTS TO DRIVE OUT ENTRANTS NEED NOT INDICATE INCUMBENT MARKET POWER—AND THE IMPLICATIONS FOR DEFINING COSTS WHERE COMMON COSTS EXIST

The premise behind the general theory for restricting reactive above-cost price cuts that drive out entrants is that such price cuts undesirably protect market power. After all, pricing above cost seems to meet a standard definition of market power, given the normal premise that firms in a competitive market price at marginal cost.¹³² Relatedly, standard analysis assumes that an ability to price discriminate implies the firm must have market power.¹³³ These premises have been particularly important in the airline industry, which is the central case cited by supporters of restrictions on reactive above-cost price cuts. Regulators and enforcement agencies have assumed that each route is its own market and that airlines that run hub-and-spoke systems must be exploiting market power if they charge higher prices in routes that connect spokes to concentrated hubs than they do on other routes with similar distance and density.¹³⁴ Given these premises, if a hub airline responds to an entrant who sells on a route connecting a hub-and-spoke by lowering its price on that route and driving out that entrant, this must reflect an undesirable protection of incumbent airline market power.

These premises have reinforced intellectual frustrations about the seeming failure to realize the predictions of contestable market theory in the airline industry. Contestable market theory held that in markets where entry was very easy it would not matter whether an incumbent firm had a one-hundred percent market share. The threat of entry alone would make the incumbent lower prices to competitive levels. Individual airline routes were

132. See 1992 Horizontal Merger Guidelines, 57 Fed. Reg. 41,552 §§ 1.0-1.22 (Sept. 10, 1992); PHILLIP AREEDA & LOUIS KAPLOW, *ANTITRUST ANALYSIS* 556 (5th ed. 1997); CARLTON & PERLOFF, *supra* note 19, at 92; JEAN TIROLE, *THE THEORY OF INDUSTRIAL ORGANIZATION* 284 (1988); DON E. WALDMAN & ELIZABETH J. JENSEN, *INDUSTRIAL ORGANIZATION* 40, 437, 667 (2d ed. 2001); William M. Landes & Richard A. Posner, *Market Power in Antitrust Cases*, 94 HARV. L. REV. 937, 939 (1981).

133. See AREEDA & KAPLOW, *supra* note 132, at 437; CARLTON & PERLOFF, *supra* note 19, at 277; WALDMAN & JENSEN, *supra* note 132, at 436; Richard A. Posner, *Oligopoly and the Antitrust Laws: A Suggested Approach*, 21 STAN. L. REV. 1562, 1578-79 (1969); Hal R. Varian, *Price Discrimination*, in 1 HANDBOOK OF INDUSTRIAL ORGANIZATION 599 (Richard Schmalensee & Robert D. Willig eds., 1989).

134. Enforcement Policy Regarding Unfair Exclusionary Conduct in the Air Transportation Industry, 63 Fed. Reg. 17,920 (proposed Apr. 10, 1998); U.S. Appellate Brief, *supra* note 36, at 5-7, 64; U.S. Summary Judgment Memo, *supra* note 36, at 40.

considered the classic example of a contestable market because it was so easy to move or lease planes to enter a route if an incumbent monopolist charged prices that were too high. Thus, the expected result of airline deregulation was that each route would be priced at competitive rates (no higher than the cost of the most efficient potential entrant) no matter how much any individual carrier dominated sales on that route.¹³⁵ When instead hub prices turned out to be persistently higher, one intellectual hangover was the suspicion that contestable market theory was not working because airlines were engaged in some anticompetitive conduct to exclude entrants. Reactive price cuts seemed one promising target. Restricting them, and encouraging inefficient entry, had the hope of forcing airlines at least to engage in a type of limit pricing that amounted to restoring the market to a contestable state. Edlin's piece is clearly in this spirit, explicitly hoping that banning reactive price cuts will make the market more "contestable."¹³⁶

But the problem is that these underlying premises were never true. In a hub-and-spoke airline system, as Section A shows, each route has interlinked demand and common costs, and thus cannot be assumed to be in separate markets in which incumbents enjoy market power. Nor does the existence of price discrimination, inside or outside of the airline industry, alone prove market power, whether one defines that as an ability to price above the "competitive" level or (more helpfully) as a power to restrict total market output in order to increase revenue or profits.¹³⁷ To the contrary, as

135. See generally Michael E. Levine, *Airline Competition in Deregulated Markets*, 4 YALE J. ON REG. 393, 395, 400-01, 403-05 (1987).

136. See Edlin, *supra* note 6, at 989-90.

137. Both definitions of "market power" are commonly used by courts and scholars. See Gregory J. Werden, *Demand Elasticities in Antitrust Analysis*, 66 ANTITRUST L.J. 363, 370, 373-74 & n.36 (1998) (collecting sources); see also Benjamin Klein, *Market Power in Antitrust: Economic Analysis After Kodak*, 3 SUP. CT. ECON. REV. 43, 44, 71-85, 88-92 (1993) (arguing for the latter definition). If a competitive market would price discriminate, then a firm that price discriminates does not price above the competitive level. Likewise, if such competitive price discrimination maximizes output by enhancing the ability of firms to incur recurring common costs, then it does not involve the exercise of any power to restrict market output in order to raise prices and revenue. Of course, one could instead define "market power" as the ability to price (to any customer) above marginal costs. See, e.g., CARLTON & PERLOFF, *supra* note 19, at 92; TIROLE, *supra* note 132, at 284; WALDMAN & JENSEN, *supra* note 132, at 40, 437, 667. And in markets where competitive firms do not use price discrimination to cover common costs, this alternative definition works fine. But in the present context, the problem is that this alternative definition presupposes just what we are investigating: whether competition forces firms to price at cost. This definition would also be misleading both linguistically and functionally because (1) it would define "market power" to exist even when the firm has no power over market output or price, but rather is one of many competitive rivals with very small market shares, and (2) it would mean that "market power" no longer indicated the ability to impose an inefficient result harmful to consumers. One could nonetheless stick to the alternative definition with the qualifier that firms often desirably exercise "market power" in highly competitive markets, but instead it is more helpful to use a definition that better tracks function and ordinary understandings. See Klein, *supra*, at 71-84, 88-92 (agreeing that defining market power in terms of an ability to price above marginal cost is unhelpful). Luckily, courts and enforcement agencies do not appear to have defined market power as an ability to price above cost. See *id.*; Werden, *supra*, at 370, 373-74 &

Section B explains, in normal (not unrealistically “perfect”) highly competitive markets where firms face common costs in delivering goods to buyers with different demand elasticities, price discrimination among those buyers is frequent and predictable.¹³⁸ Such price discrimination charges more to the high-demand buyers, thus effectively recouping a higher share of common costs from them. While competition will drive total economic profits to zero, and thus make total revenue equal total costs, competition will also force each firm to adopt, where sustainable, the price-discrimination schedule that maximizes its profits.¹³⁹ The competitive result thus may be a price-discrimination schedule that, by maximizing the revenue earned from any common costs, also maximizes each firm’s ability to incur common costs, and thus maximizes industry output and aggregate sales to the high- and low-demand buyers.

But this competitive result will not always be stable because firms or entrants will constantly be tempted to deviate from the output-maximizing price-discrimination schedule by serving only (or mainly) the high-paying buyers at a lower price. If they do so, Section C points out, other firms will have to follow suit by cutting their prices to these high-paying customers, given that retaining them is necessary to cover common costs. Because this deviation is inefficient, lower industry output will result as long as it lasts. If the entrant’s costs of serving only the high-paying customers are higher than the costs of serving those customers that would be incurred by the firms that also serve lower-paying customers, then separate provision is in fact not as efficient as common provision. The firms that serve both sets of buyers will thus be able to drive the entrant out with a price that is above their costs of serving the high-demand buyers, and after they have done so, raise prices to the high-paying buyers and restore the output-maximizing price-discrimination schedule. That is, reactive above-cost price cuts that drive out entrants not only do not necessarily signal the undesirable protection of market power, they may be the normal and necessary way of restoring efficient price discrimination in competitive markets.

n.36. In any event, whatever definitional labels we choose will not alter the substantive result that what is being described is a desirable state of affairs, and that unilateral pricing decisions that protect it should thus not be condemned by antitrust law.

138. See Michael E. Levine, *Price Discrimination Without Market Power*, 19 YALE J. ON REG. 1, 6-7 (2002); William J. Baumol, *Normal and Effectively Competitive Equilibrium with Ubiquitous Discriminatory Price Taking* (Apr. 26, 2002) (unpublished manuscript, on file with author); see also Klein, *supra* note 137, at 65-66, 71-72 (describing this phenomenon without linking it to the need to cover common costs).

139. Baumol, *supra* note 138 (manuscript at 1-6).

A. *Individual Routes in Hub-and-Spoke Systems Cannot Be Assumed To Be Separate Markets*

The premise that airline markets were properly defined by a route between city *A* and city *B* failed to recognize that the advent of hub-and-spoke systems of airline travel makes it problematic to separate individual routes from a general network of airline flights. With a hub-and-spoke system, airlines can satisfy customers who desire travel between a multitude of city-pairs with dramatically fewer flights and less cost by having one “hub” city with flights to each of the other “spoke” cities.¹⁴⁰ Moreover, the flights will be fuller (and thus cheaper per passenger) in the hub-and-spoke system, and more likely to sustain a reasonable schedule of travel on the larger jet planes that passengers prefer because of their more comfortable ride.¹⁴¹ Indeed, it is clear that without hub-and-spoke systems it would not be possible to sustain a reasonable schedule of air travel from small cities that may have hundreds of people traveling *somewhere* each day (who could thus fill a flight to a hub), but only a handful of people traveling to any single city (who could thus not cover the cost of a schedule of nonstop flights from their small city to all their separate destinations).¹⁴²

The efficiencies driving this hub-and-spoke system are thus overwhelming. But these efficiencies mean one cannot simply assume routes between different cities are separate markets. Passengers with different itineraries are being combined on the same flights. The market prices for seats on a flight from hub city *A* to spoke city *B* thus turn not just on the demand for travel between those cities, but also on the demand for travel between city *B* and cities *C-Z*. And travel between city *B* and cities *C-Z* might be serviced by rivals not only through the same hub, but with nonstop flights or through other hub cities. Some passengers might be interested only in nonstop flights between a hub and a spoke city, but in a sense they are side beneficiaries of a system driven mainly by the need to provide hub-and-spoke coverage. Indeed, the hub-and-spoke system makes possible nonstop service between cities that otherwise would not be possible.¹⁴³ With the routes intermingled in this way, it may thus make much more sense to think about the entire hub-and-spoke network as the relevant product an airline provides.¹⁴⁴ If the hub-and-spoke network itself is the relevant product, then the relevant price and cost would be those

140. See Levine, *supra* note 135, at 441-46.

141. *Id.* at 441-42.

142. *Id.* at 442-43.

143. *Id.* at 443.

144. See generally 1992 Horizontal Merger Guidelines, 57 Fed. Reg. 41,552 §§ 1.0-1.22 (Sept. 10, 1992) (examining buyer and supplier substitution to define antitrust markets).

earned and expended across the hub-and-spoke network, not on individual routes.

Even if individual routes are separate markets for some purposes, their integration into a hub-and-spoke system requires incurring large common costs, whose allocation across the constituent routes is largely arbitrary.¹⁴⁵ Developing such a hub-and-spoke system requires large investments. The airline cannot just have flights between the most attractive city-pairs to reap the advantages of hub-and-spoke travel but must rather have a full network of flights. It must have sufficient gate slots and ticketing offices; a fleet of planes and equivalent maintenance facilities; baggage transfer operations; a large team of trained personnel; and a complex system for marketing, planning, scheduling, reserving, dispatching, and pricing across the entire hub-and-spoke network. Perhaps more important, it must incur the costs of maintaining flights that impact revenue on connecting flights, and incurring increased ground time for planes in order to provide connections, which also generally entails using more gates.¹⁴⁶ Moreover, in order to maintain the reliability of its hub-and-spoke system over time, the airline probably has to commit to covering certain routes even though they become unprofitable over the short run.

Consideration of hub-and-spoke economics thus sharply undercuts the intuition that reactive price cuts that drive out entrants on a particular airline route undesirably protect market power on that route. In a hub-and-spoke system, individual routes may not represent separate markets because of interlinked demand and common costs. If so, then airlines should not be considered to have market power unless they can constrain output on the entire hub-and-spoke system to raise revenue or profits. But in fact, airlines do not earn above normal profits, suggesting that they probably do not have such market power.¹⁴⁷ Even if the price on a single route falls below the

145. See Baumol, *supra* note 41, at 59 (noting that any allocation of common costs is arbitrary). Airlines sometimes allocate these common costs by simply dividing the total hub-and-spoke costs by the number of flights or flight hours. *United States v. AMR Corp.*, 140 F. Supp. 2d 1141, 1175-77 (D. Kan. 2001). But while this may make sense for accounting or business purposes, as an economic matter, any allocation of joint production costs is inherently arbitrary. For example, more than fifty percent of passengers in major hubs are "connecting passengers" (they are flying through the hub between two spokes), see Edlin, *supra* note 6, at 944 n.12, which leaves something less than fifty percent as "hub passengers" (the hub is one end point of their travel). Thus, one could take the view that because the connecting passengers would support the relevant flight, the incremental cost of flying the hub passengers is extremely low. See *id.* Alternatively, one could take the view that because the hub passengers would support the flight, the incremental cost of flying the connecting passengers is extremely low. The problem is that both could be true, making any allocation of joint costs to the individual flights arbitrary.

146. An airline that does not offer connecting flights (like Southwest Airlines) has very little turnaround time because it needs just enough time to unload one set of passengers and load the next. It need not keep the plane waiting for connecting customers. This shorter ground time also means it can run more flights per gate, reducing the capital costs of planes and gates.

147. See Baumol, *supra* note 138 (manuscript at 8); Gary J. Dorman & William J. Baumol, *On Cures That Bring Their Own Diseases* 4 (Apr. 14, 2002) (unpublished manuscript, on file with

separate cost of flying that route, continuing that price can be efficient if it increases demand on other routes in the hub-and-spoke system. In that case, under a cost-based test, the prices a hub-and-spoke airline charges should not be considered predatory unless the overall revenue on a hub-and-spoke system falls below the cost of providing the entire hub-and-spoke system.¹⁴⁸

If demand for each route could be separated, the existence of widespread common costs still means that prices should not be considered predatory under a cost-based test unless either (a) prices across the system are lower than system costs or (b) the price on the particular route falls below a measure of separate costs that excludes all common costs of operating the hub-and-spoke system. This is why the district court was correct to reject predation tests that compared individual route prices to fully allocated system costs in the American Airlines litigation.¹⁴⁹ An analogy might be drawn to the pricing used to recoup the common costs of flying a plane. As everyone who travels knows, some seats are sold for much more than others on the same flight, which may well mean that the lowest prices charged are well below the average cost per seat. But given the dominance of common costs, this fact should not make prices on those seats predatory. Instead, one must either compare the incremental revenue for the flight to the incremental cost of making that flight,¹⁵⁰ or compare the price on the lowest priced seats to the incremental cost of serving the additional passenger, which may well be extremely low since costs are almost the same whether the seat is empty or full.¹⁵¹

This is not a proposition unique to airlines, but rather is just one instance of the more general proposition that products with common costs should be considered below cost only if the price for any one product is lower than its separate cost (which is unlikely since that excludes common costs) *or* if the price for the combination of products falls below their combined cost (which includes common costs).¹⁵² Likewise, if one product with common costs is sold at different prices to different sets of customers,

author). Such evidence is suggestive but not necessarily determinative. A firm could offset supracompetitive profits on some routes with greater inefficiency on other routes. Or perhaps airlines earn supracompetitive profits that they distribute to their unions. But one would think that the inefficiency on other routes would be driven out in the long run. Also union power in other industries does not generally eliminate supracompetitive profits, and it is difficult to explain why airline unions would drive so many airlines to bankruptcy if they enjoyed supracompetitive profits.

148. If the conclusions of Part II are accepted, the relevant costs would be the costs that are varied by the predatory increase in output. That also is true for all references to separate or common costs in this Section.

149. See *supra* Section I.A.

150. See *Int'l Travel Arrangers v. NWA, Inc.*, 991 F.2d 1389, 1396 (9th Cir. 1993).

151. See 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 740b2, at 427. In addition, one would have to show that the predatory output increase reflected in these seats sufficed to drive out the relevant rival output. See *supra* Section II.C.

152. See Baumol, *supra* note 41, at 59-61.

the prices should be deemed below cost if the price to any one set of customers is below the separate cost of producing that quantity *or* if the prices recovered from the combination of customers is lower than the combined costs of producing the aggregate quantity.¹⁵³ Indeed, if a multiproduct firm cuts prices on one product (or to one set of customers), but its prices for the combination of products (or customers) still cover all costs (including common costs), the initial prices for the combination of products (or customers) must have exceeded their combined cost and been supracompetitive. Thus, the price cut on one product (or to one set of customers) without a corresponding increase on other products (or customers) amounts to a desirable discount from oligopoly or monopoly prices.¹⁵⁴

Yet this does not mean that the existence of price discrimination across routes shows airlines must have market power, or that all reactive above-cost price cuts undesirably protect airline market power, for reasons considered next.

B. *Why Competitive Markets May Induce Price Discrimination That Maximizes Output*

Airlines recoup the common costs of a hub-and-spoke system not with uniform prices, but with a complex regime of prices that vary sharply not just from route to route, but from customer to customer and day to day. Since passengers flying nonstop between a hub-and-spoke city get a more valuable slice of the hub-and-spoke system (quicker, more convenient travel), they are charged more per mile.¹⁵⁵ Indeed, passengers on any single flight are charged wildly different prices to recover the common costs of operating that flight. Such price discrimination is not a feature unique to the airline industry. It also occurs with movie theaters that (without any market power) cover the common costs of exhibiting movies by price discriminating among adults, seniors, and children, or with retailers that (without any market power) cover the common costs of operating retail space by charging different markups on different goods with the same low

153. See *id.* at 63-65; cf. 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 724d, at 292-93, ¶ 742c, at 460-61, ¶ 742c2-d, at 464-68 (reaching similar conclusions, though sometimes, for unclear reasons, requiring proof of both rather than either).

154. Cf. 10 PHILLIP E. AREEDA, EINER ELHAUGE & HERBERT HOVENKAMP, *ANTITRUST LAW* ¶ 1758f (1996) (establishing a similar proposition for package discounts offered in tying cases).

155. This is true even though the very features that make their air travel more valuable also might make it seem that their travel is less costly. They take a more efficient route, require fewer takeoffs and landings, need no arrangements to make sure connections are made, and do not require multiple sets of baggage handling. Which passengers enjoy these advantages of directness, however, is itself a product of how the hub-and-spoke system is structured. Moreover, as noted above, the allocation of joint costs is inherently arbitrary.

marginal retailing cost.¹⁵⁶ Price discrimination among buyers who can be served only by incurring common costs is thus routine even in highly competitive markets, including hotels, computers, automobiles, books, clothing, groceries, restaurants, telecommunications, and the vast range of other products that offer coupons, rebates, student or senior discounts, quantity discounts, or different prices at different times or places.¹⁵⁷ Indeed, it is hard to think of industries without price discrimination, even though most of these industries are highly competitive or contestable, and the firms in them earn zero economic profit (i.e., a normal rate of return).

The prices charged to buyers thus vary greatly in competitive markets where, as typical, common costs exist. At a minimum, these prices will cover the separate costs of serving each customer—that is, the additional cost imposed by a single customer assuming common costs have been incurred. But if that were the only price charged to all customers, then firms could not recover their common costs and thus would never incur them. Accordingly, the price schedule for serving the combination of buyers will have to cover the common costs of doing so. Where price discrimination is sustainable, competition will force firms to adopt the price-discrimination schedule that maximizes the revenue from customers for any common costs that have been incurred.¹⁵⁸ But competition will also assure that the total revenue earned from both sets of buyers will not exceed total separate and common costs, and thus economic profits will remain at zero even though some buyers are paying a price above the separate cost of serving them.¹⁵⁹ That is, a price-discrimination schedule that earns positive economic profits will be undercut by a rival or entrant in a competitive or contestable market. But a price-discrimination schedule that maximizes the revenue from any common costs without exceeding them cannot be undercut by a profitable schedule of price discrimination that serves both sets of customers.

Where sustainable, this competitive price discrimination will generally increase output. If all buyers were charged the same price, and thus effectively covered both their separate cost and an equal share of common costs, then sales would be lost to those buyers who are not willing to pay that high a price but would be willing to pay a price higher than their separate costs. This would mean an inefficient reduction in output since

156. Sometimes, as with movie seats, the products sold are identical but firms differentiate among buyers who have different demand elasticities. Other times, product differences are created that may even degrade some output in order to help differentiate among buyers, as with airline flight restrictions, cars, computer equipment, and other products and services. See Klein, *supra* note 137, at 65-66; Levine, *supra* note 138, at 20-21, 23-27; Baumol, *supra* note 138 (manuscript at 15-16).

157. See Baumol, *supra* note 138 (manuscript at 1, 6); Klein, *supra* note 137, at 65-66; Levine, *supra* note 138, at 2-3, 14-16, 21-29.

158. Baumol, *supra* note 138 (manuscript at 3-6).

159. *Id.*

their willingness to pay exceeds the marginal cost of doing so if common costs can be covered. If instead the buyers who value the product more highly can be charged a higher price, and thus cover a higher share of common costs, then the buyers who value the product less highly can be charged a lower price that does not cover their proportionate share of common costs but remains above their separate costs. Additional sales can be made and output will be expanded. In other words, the price-discrimination schedule that maximizes the revenue from any common costs that are incurred will also enable firms to incur the most common costs and maximize industry output.

The tricky issue is explaining how such discriminatory prices can be sustained without market power given the incentives of individual firms to increase profits by concentrating their sales on high-demand buyers. Prior literature has established at least two situations where such desirable price discrimination will occur without any market power to reduce market output in order to reap supracompetitive profits.

First, there might be economies of scale, so that some common costs might suffice to serve the entire market (e.g., one plane suffices to serve the route), and yet the market might be contestable in that any entrant could costlessly enter with similar economies of scale. If so, then discriminatory Ramsey pricing (which charges more to less elastic buyers) will prevent any competitive entry and thus be sustainable.¹⁶⁰ Indeed, that threat of entry will drive the firm to discriminatory pricing.¹⁶¹ Since the market can efficiently sustain only one firm, no rival can hope to retain a disproportionate share of high-demand buyers in the market with less discriminatory pricing. Yet the competitive threat of entry prevents any incumbent from cutting market output to raise revenue or from pricing in ways that make revenue exceed total costs.

This can also describe workable competition in markets where each firm serves some limited set of customers exclusively (like the local corner store) but would, if any firm's total revenue exceeded costs, provoke rivals into moving in and serving that set too. In these cases, rivals cannot concentrate sales on the high-demand buyers because local economies of scale effectively bundle them with low-demand buyers. One can reach similar results with spatial models of brand preferences, where each brand's characteristics have a particular "location" that matches the preferences of a set of high- and low-demand buyers but that cannot earn supracompetitive profits because, if it did, other firms would create a brand with the same characteristics. In effect, each brand can be a contestable market. If so, then

160. See WILLIAM J. BAUMOL ET AL., *CONTESTABLE MARKETS AND THE THEORY OF INDUSTRY STRUCTURE* 208-17 (1982).

161. Baumol, *supra* note 138 (manuscript at 1-6).

each brand can and will engage in discriminatory pricing to cover its common costs (including the costs of brand advertising) even though it competes in a larger market and cannot raise its price-discrimination schedule to a level that reaps supracompetitive profits.¹⁶²

Second, individual buyers may be high-demand buyers for some products and low-demand buyers for other products that are efficient to buy from the same seller. For example, restaurant patrons buy food and drink at the same meal, adult moviegoers often buy tickets for themselves at the same time that they buy tickets for their kids, and consumers often prefer buying multiple products after making a single trip to one retailer (like a supermarket).¹⁶³ If purchases are efficiently bundled in this manner, it has been shown that competitive sellers can maximize buyer utility by offering a discriminatory schedule of Ramsey prices that charges higher markups on the products for which buyers have less elastic demand, producing total revenue that covers the common costs of operating the retail space but does not exceed total costs.¹⁶⁴ Rivals cannot undercut this with a less discriminatory price schedule, for they will be offering less utility to buyers and are thus less likely to attract buyers into making the trip to their place of business. This probably explains why, for example, competitive restaurants (from fast food joints to fancy restaurants) charge much higher markups on drinks than they do on food even though the latter is more costly and difficult to provide. Since in a competitive market the total revenue on food and drink must match total cost, no rival can lower prices just on the high markup drinks without raising them on food. And since the bulk of patrons consume both food and drink in the same seating, a rival that offers lower drink markups with compensating higher food markups cannot normally hope to get a disproportionate share of drinkers. Rather, because the rival has undercut the price schedule that maximizes patron utility, it will attract fewer patrons.

We might extend this second theory beyond the retail context to the general case where brand loyalty causes individuals to commit to one brand for multiple products until another brand offers a more attractive package. In the retail context, it is efficient to bundle the purchase of different

162. See Severin Borenstein, *Price Discrimination in Free-Entry Markets*, 16 RAND J. ECON. 380, 380-81 (1985); see also Klein, *supra* note 137, at 72, 77-78 (arguing that almost every brand faces a downward sloping demand curve that permits price discrimination).

163. This would also apply to purchases from hotels or integrated packages like computer systems. See Christopher Bliss, *A Theory of Retail Pricing*, 36 J. INDUS. ECON. 375, 391 (1988).

164. See Mark Armstrong & John Vickers, *Competitive Price Discrimination*, 32 RAND J. ECON. 579, 580-82 (2001); Bliss, *supra* note 163, at 378-80, 382, 385. One could see this as a special case of a contestable market. "A shop enjoys a limited but significant natural monopoly over the demand of the shopper who has incurred the cost of coming to the shop." Bliss, *supra* note 163, at 378. If retail shops try to earn supracompetitive profits, however, it is easy for rivals to "enter" this market by persuading customers in the future to incur the costs of traveling to their shop instead.

products either because buyers enjoy them more together (food and drink, sitting with one's kids at the movies) or because assembling goods in one place saves consumer-transportation and -time costs.¹⁶⁵ For brands, one might similarly say that combining goods under one brand name saves consumer-information or -search costs because positive experience with one brand can be applied to the next. One might understand various airline efforts to build brand loyalty, including frequent flier programs, as partly an effort to help maintain the proportion of business and personal travelers by shifting individual consumers (who take both business and personal flights) from making flight-by-flight choices to instead choosing between airlines for their full range of flight needs. To the extent such efforts are successful, competing airlines would have incentives to engage in intrapersonal price discrimination by offering consumers the discriminatory price schedule for business and personal travel that maximizes their utility. In these cases, rivals cannot concentrate sales on the high-demand buyers because their high-demand purchases are bundled with their low-demand purchases.

Although these two situations explain some of the phenomenon, the prior literature has not yet explained successfully just how competitive markets can arrive at and sustain efficient output-maximizing discriminatory pricing in markets where high- and low-demand purchases are not bundled and multiple firms compete for the business of the same buyers. I thus here offer two additional models to explain how such competitive price discrimination can be maintained.

First, suppose the buyers willing to pay a low price are informed and price-sensitive, while the buyers willing to pay a high price are uninformed or price-insensitive (below some level) and select sellers at random. Being uninformed really means being unwilling to incur the costs of becoming informed about price differences. This may reflect the fact that spending time is more costly for these buyers, or that a given information cost matters less to them because of their relative price insensitivity. For example, compared to tourists, business travelers might well be less price-sensitive because they are not spending their own money or find it too costly to spend the time investigating prices. If so, price discrimination between them can be maintained, because a rival or entrant who undercuts the high price will not gain a greater-than-random share of business travelers.

Further, it seems reasonable to assume that the random choices made by uninformed (or price-insensitive) buyers are not so much among firms as among flights or retail outlets. That is, the buyer uninformed about (or insensitive to) relative prices just buys whatever flight is most convenient or from whichever retail outlet the buyer happens to enter. If so, a firm that

165. See Bliss, *supra* note 163, at 375, 377-78.

incurs more increments of common costs by offering more flights or retail outlets gains a greater share of the high-demand buyers. Since gaining a greater share of these high-demand buyers is more profitable, each firm will have incentives to compete by increasing the common costs it incurs. Further, since these high-demand buyers are willing to pay more than the average cost of fully utilized common costs, firms will be willing to expand their increments of common costs (like number of flights) to the point where they have some unused capacity (like empty seats). And rather than let that unused capacity go to waste, they might as well sell it to the low-demand (informed) buyers at a price closer to their marginal costs once common costs are incurred. This additional revenue will lead to additional expansion until competitive firms reach the maximum output allowed by the use of discriminatory pricing to cover common costs. Firms whose discriminatory price schedule produces revenue that exceeds common costs will be undercut by a schedule that does not; likewise, firms that have common costs that are inefficiently high can be undercut by an efficient firm that will incur lower common costs and offer a lower price schedule at every level.

Edlin offers a related model where (in a competitive market with low entry barriers) retail outlets price discriminate among informed and uninformed buyers with price-matching policies.¹⁶⁶ He reaches the quite different conclusion that it will lead all firms to adopt a supracompetitive price (with an offer to match a lower price that is never met) that will keep inducing entry by firms that adopt the same pricing strategy, leaving each retail store with fewer uninformed buyers, until per-unit costs rise enough to dissipate supracompetitive profits fully.¹⁶⁷ But if his prediction actually occurred, then each firm would have unused capacity that would create powerful incentives for it to price discriminate—offering that unused capacity at lower prices to the informed marginal buyers who will not buy at the supracompetitive price. The fact that other firms will match that price to informed buyers will not deter this because selling this unused capacity at a lower price to some share of the informed buyers is more profitable than letting it go to waste. Indeed, such price matching by other firms would reinforce the conclusion that such price discrimination would spread

166. See Aaron S. Edlin, *Do Guaranteed-Low-Price Policies Guarantee High Prices, and Can Antitrust Rise to the Challenge?*, 111 HARV. L. REV. 528, 529-31, 536-52, 573-75 (1997). Professors Salop and Stiglitz earlier offered a related model whereby uninformed buyers choose shops at random and informed buyers do not, and proved that one possible market equilibrium was price dispersion with some shops at high prices and other shops at lower prices. See Steven Salop & Joseph Stiglitz, *Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion*, 44 REV. ECON. STUD. 493, 494, 502-07 (1977). But they assume each shop offers only a single price, and thus do not consider the possibility that each shop might price discriminate among informed and uninformed buyers.

167. Edlin, *supra* note 166, at 542-43, 547-49.

across a competitive market. Alternatively, if existing firms have grown so inefficient that their marginal costs have risen to match their supracompetitive prices, then an entrant would have incentives to enter (without adopting similar inefficiencies) with the same high price for uninformed buyers but a lower price for informed buyers that incumbents (given their high costs) would no longer be able to match, thus giving the entrant all the informed buyers unwilling to pay the supracompetitive price. Either way, competition and free entry is not consistent, as Edlin supposes, with firms operating in such a way that their total revenue exceeds the most efficient cost level.

Second, a different model can also explain desirable output-maximizing competitive price discrimination even among informed buyers who can choose among multiple firms and do not make high- and low-demand purchases from the same firm. We need only posit three conditions (in addition to the general assumption that firms can distinguish high- and low-demand buyers and prevent them from reselling to each other). (1) At the same price, the share of high-demand customers each firm receives depends on how many increments of common costs it incurs. This seems a reasonable assumption. The more flights or movies or retail locations a firm offers, the greater the share of customers it should get at any equal price. This assumption has the crucial feature that it means that (at the same price) firms compete for high-demand customers by incurring more common costs. (2) There is a discontinuity of demand that results in a set of high-demand customers willing to pay more than the sum of separate costs and an equal share of common costs, and low-demand customers who are willing to pay less than that but something above their separate costs. (3) The market is sufficiently competitive that supracompetitive profits (revenue that exceeds costs) invite rapid expansion or entry, and any price cut to high-demand customers is rapidly matched.¹⁶⁸

With those assumptions, one can show that firms on a competitive market will have incentives to price discriminate in ways that expand output. This is established in a mathematical footnote,¹⁶⁹ but a concrete

168. The importance and justification for the rapid price matching assumption is addressed *infra* Section III.C.

169. Define the following variables. C_s means the separate costs incurred per buyer. C_c means the common costs necessary to serve up to X buyers (e.g., for a flight with X seats). H is the total number of high-demand passengers willing to pay up to $C_s + C_c/X + S$, where S is whatever sum they are willing to pay above their proportionate share of common costs. L equals the total number of high-demand passengers willing to pay more than C_s but only up to $C_s + C_c/X - T$, where T is whatever amount less than their proportionate share of common costs they are willing to pay. The condition that they are willing to pay more than C_s means $C_c/X > T$. P_H means the price charged to high-demand customers and P_L means the price charged to low-demand customers. The market is assumed to be sufficiently competitive that any P_H charged by one firm that undercuts the other firms is immediately matched by them. *See infra* Section III.C. N_i means the number of incremental common costs (like flights) incurred by the firm in question with a total of I firms,

example may help. Suppose the common costs of flying a plane with 200 seats are \$30,000 and the separate costs incurred for each additional passenger are \$50. A full plane thus costs \$40,000 to fly, or \$200 per passenger. A half-full plane costs \$35,000 to fly, or \$350 per passenger. Suppose there are a total of 1000 possible passengers: 500 business travelers who will pay up to \$300 per flight and 500 tourists who will pay only up to \$100 per flight. Suppose further that there are two airlines that exhibit competitive behavior, instantly matching each other's prices and that (at those equal prices) passengers divide up between the two firms in proportion to the number of flights they offer. If the airlines charged a uniform price of \$200 to cover average common costs, then the 500 tourists will not fly. The 500 business travelers will want to fly, and will fill one flight for each airline. But neither airline (nor an entrant) can afford to fly

and N means the total number of incremental common costs incurred by all firms in a competitive market.

Since each firm gets high-demand buyers in proportion to the common costs it incurs (e.g., flights it offers), the number of high-demand buyers a firm will enjoy is H^*N_i/N . Let us call the number of low-demand customers that firm enjoys L_i . A firm will incur common costs of N_i only if $P_H^*H^*N_i/N + P_L^*(L_i) \geq N_i^*C_c + C_s^*(L_i + H^*N_i/N)$. Suppose each firm charges a uniform price that just covers common costs at full capacity, i.e., a price equal to $C_c + C_c/X$. No low-demand customers will buy at this price. Thus, L_i for each firm, and L for all firms, will equal 0. Plugging this into the above equations, we get that common costs of N_i will be incurred only if $(C_c + C_c/X)^*H^*N_i/N \geq N_i^*C_c + C_s^*H^*N_i/N$, which can be rearranged as $H \geq X^*N$. Thus, a firm will incur common costs of N_i only if total demand from high-demand customers equals or exceeds total industry capacity after N_i is added. Total capacity and output will be the largest N^*X that is less than H . Unmet demand by low-demand buyers will be L . Unmet demand from high-demand buyers will be $H - N^*X$.

This unmet demand by high-end customers should—unless capacity is added—bid up prices for sales to high-demand customers up to what these high-demand customers are willing to pay, or $C_s + C_c/X + S$. This will offer supracompetitive profits. Again, L_i will equal 0. Some firm or entrant will then be willing to add capacity in order to get a greater share of the high-demand customers by incurring common costs of N_i if $(C_s + C_c/X + S)^*H^*N_i/N \geq N_i^*C_c + C_s^*H^*N_i/N$. This is the same as $H + H^*X^*S/C_c \geq N^*X$. Thus, firms will be willing to add capacity up until the point that $H + H^*X^*S/C_c \geq N^*X$. This will mean excess capacity (empty seats) of $N^*X - H$, which is up to H^*X^*S/C_c . (Even if the price to the high-demand customers is bid down somewhat, any price that exceeds the sum of separate costs and an equal share of common costs leads to the same sorts of results. And any price that does not exceed that sum will be bid up since demand will exceed capacity.) Rather than allow that unused capacity (empty seats) to go to waste, each firm has incentives to offer it at a lower (discriminatory) price to the low-demand customers.

When firms do fill the other seats with discriminatory prices, then that additional profit will induce capacity to be added as long as $P_H^*H^*N_i/N + P_L^*L_i \geq N_i^*C_c + C_s^*(L_i + H^*N_i/N)$. At full discriminatory prices, this will be true when $(C_s + C_c/X + S)^*H^*N_i/N + (C_s + C_c/X - T)^*L_i \geq N_i^*C_c + C_s^*L_i + C_s^*H^*N_i/N$, which is the same as $(C_c/X + S)^*H^*N_i/N + (C_c/X - T)^*L_i \geq N_i^*C_c$. The number of low-demand buyers each firm serves will equal the capacity that would otherwise go unused, which is $N_i^*X - H^*N_i/N$. Thus, the above can be expressed as $(C_c/X + S)^*H^*N_i/N + (C_c/X - T)^*(N_i^*X - H^*N_i/N) \geq N_i^*C_c$, which can be rearranged as $H^*S/T + H \geq N^*X$. Thus, with discriminatory pricing, firms will add capacity up to the point where $H^*S/T + H \geq N^*X$. This capacity and output will be greater than with maximum uniform pricing whenever $H^*S/T + H > H + H^*X^*S/C_c$, which can be rearranged as $C_c/X > T$. And that is true (as noted above) whenever the low-demand customers are willing to pay a price above the separate costs of serving them. Thus, output will always be higher with price discrimination that charges less to low-demand customers as long as they are willing to pay something above their separate costs—that is, as long as they are willing to make some contribution toward common costs.

the other 100 on a half-full plane because the revenue (\$20,000) would be less than the cost (\$35,000). Thus, unmet demand will include not only the 500 tourists but 100 business travelers. This unmet demand will bid uniform prices up to the \$300 these passengers are willing to pay. But that would mean a per-flight revenue (\$60,000) that exceeded cost (\$40,000). Those supracompetitive profits would give each firm incentives to add one more flight each, giving each 250 total business passengers, with a per-flight revenue (\$37,500) that exceeds cost (\$36,250). Neither would add the market's fifth flight at this uniform price because that would give each flight 100 passengers and thus produce revenue (\$30,000) that was lower than cost (\$35,000). Thus, with uniform pricing, the maximum market output is four flights and 500 passengers.

But this uniform pricing results in 75 empty seats per flight, which gives each firm incentives to sell those empty seats at a lower price of \$100 to tourists. That, in turn, raises per-flight revenue to \$45,000 compared to a cost of \$40,000. Those supracompetitive profits will induce one firm (or an entrant) to add another flight at the same level of discriminatory pricing, creating a market total of five flights (each with 100 business passengers paying \$300 and 100 tourists paying \$100) with revenue-matching costs. Output will thus rise from four flights with 500 passengers without discrimination to five flights with 1000 passengers with price discrimination. Business passengers will have more flight options with price discrimination, and tourists who otherwise could not fly will be able to do so.

One might fear that one firm would drop flights to concentrate on the high-demand buyers. For example, the firm with two flights could do so figuring that flying two flights at cost is less profitable than flying one flight of 200 business travelers at \$300 and reaping \$20,000 in supracompetitive profits. But if it only offers one flight compared to the three offered by rivals, it would only get one-fourth of the business travelers at an equal price. Although this would look profitable if rivals did not respond because the revenue (\$45,000) would exceed cost (\$40,000) for all airlines, in a competitive market rivals would instantly respond by adding their own flight, meaning that the airline that dropped a flight would now only get one-fifth of the business travelers. It would thus still have revenue that matched costs but would have cut its market share in half, which should deter the move,¹⁷⁰ and even if it did not, competitive price discrimination would be restored.

All this may seem inconsistent with ordinary notions that the competitive equilibrium forces pricing at marginal cost. Indeed, theorists

170. See *infra* Section III.C (discussing the parallel issue of cuts in prices to gain a disproportionate share of high-demand buyers).

arguing that competitive markets feature price discrimination have claimed in part that such price discrimination is necessary to recoup *nonmarginal* sunk capital costs.¹⁷¹ But this claim seems dubious. If costs are truly sunk, they should and will be ignored by firms in pricing. This does not raise the paradox, as is often supposed,¹⁷² that sunk capital costs will never be incurred. If marginal costs increase with output and firms are below average costs, then it is true that pricing at marginal costs will prevent firms from incurring further sunk capital costs. But they shouldn't; rather, it is more efficient for firms to increase output with current capacity. If increasing output eventually drives marginal costs above average costs, then pricing at marginal cost will allow recovery of sunk capital costs, and will thus not prevent firms from incurring those sunk costs to expand capacity.¹⁷³

Rather, the argument that competition may require discriminatory pricing seems appropriately limited to whichever costs are truly variable over the relevant pricing period, though this may well include capital or fixed costs that are recurring over time or as output rises. However, the existence of recurring common costs means the marginal cost curve takes a different shape than the conventional assumption that it continually slopes upward. Where there are common costs to serving a set of customers, the costs of incremental increases in output are lumpy and discontinuous, featuring a large cost when the common cost is incurred (of adding a flight or showing a movie) followed by much lower separate costs (of seating another customer), then perhaps another large incremental common cost (if another flight or movie is added) followed by lower separate costs, and so on. Further, the dichotomy in demand means some buyers are above and some below the average per-person cost of incurring the common cost at full capacity. Thus, rather than the traditional graph, such a situation may best be reflected in something like Figure 1.¹⁷⁴

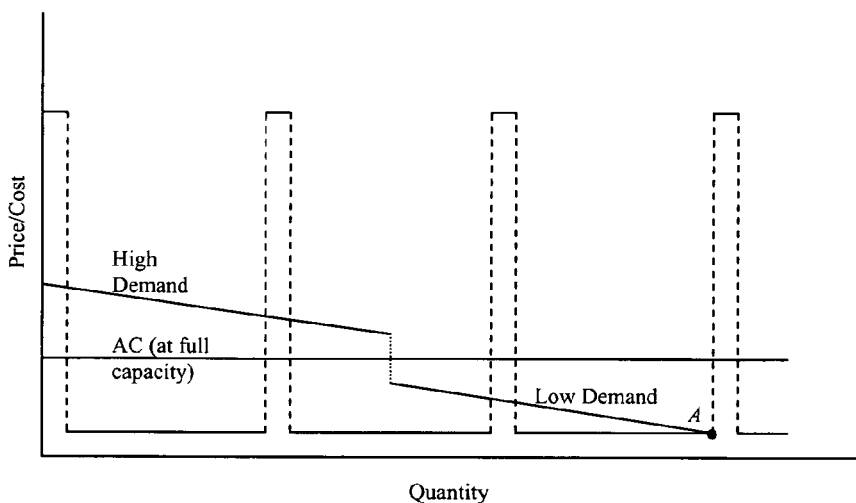
171. See Klein, *supra* note 137, at 90; Levine, *supra* note 138, at 8 n.21, 11-12, 17; Baumol, *supra* note 138 (manuscript at 10).

172. See, e.g., Levine, *supra* note 138, at 11-12.

173. If instead marginal costs decline across industry output, then we have a situation of natural monopoly. The firm that first incurs those sunk capital costs will become a monopolist, and monopoly returns will provide ample incentive to incur those sunk costs. Subsequent firms will not incur sunk capital costs, nor should they—because when the market is a natural monopoly it would be wasteful for them to do so. To be sure, any government rate regulation to limit those returns must be constructed to allow recovery of sunk costs. And the most efficient way of doing so is to impose Ramsey pricing, which is a price-discrimination schedule that prevents monopoly profits but charges high-demand buyers more than low-demand buyers so that the higher-demand buyers cover a greater share of the sunk costs but the lowest-demand buyers pay prices closer to the low marginal cost of the final units of production. See William J. Baumol, *Ramsey Pricing*, in 4 THE NEW PALGRAVE: A DICTIONARY OF ECONOMICS 49, 49-51 (John Eatwell et al. eds., 1987). But such government imposition of a discriminatory price schedule is in no sense *necessary* for the recovery of truly sunk costs.

174. These incremental common costs are for graphical purposes assumed to be constant, but if (as typical) industry costs generally increase with output, then it is more likely that as output increases each common cost spike will be somewhat higher than the spike that preceded it.

FIGURE 1.



If firms tried to price uniformly at the level (*A*) where the demand curve intersected the marginal cost curve, then they would lose money because they could not recover the variable common costs reflected in the first three cost spikes.¹⁷⁵ They thus would not incur those common costs at that price. Note also that no single consumer has marginal demand that comes close to the full marginal cost of taking on any incremental common cost. Rather those common costs must always be allocated among some set of buyers whose marginal demand is lower but who in combination have enough demand to justify incurring the common costs. With uniform pricing, one could try to allocate just those common costs among the high-demand buyers reflected in the leftward portion of the demand curve. But that would fail to include some buyers whose marginal demand does exceed marginal costs once common costs are incurred, and thus result in lower output. With discriminatory pricing, firms instead allocate some of the high- and low-demand buyers to each increment of common cost, thus covering that increment of common cost mainly with the high-demand buyers but also selling to the low-demand consumers at prices down to marginal cost, and efficiently expanding output up to point *A*. Accordingly, such competitive price discrimination does not mean a deviation from the normal rule that firms will price at marginal cost—it rather reflects the

175. The demand curve does not really cross the cost curve where it passes through the dotted lines, because they reflect discontinuities in costs, so that the demand curve is always above or below the marginal cost curve until it reaches point *A*.

mechanism by which, given common costs, firms are able to achieve marginal prices that come as close as feasible to marginal cost.

In short, price discrimination does not prove market power. Thus, no market power is proven by the airline practice of charging higher prices on routes that connect spokes to concentrated hubs than on other routes.¹⁷⁶ Indeed, empirical evidence shows that price dispersion increases with greater airline competition.¹⁷⁷ Instead, it may not make sense to conclude an airline has market power unless it dominates and earns monopoly profits on the whole set of cities connected by a hub-and-spoke system, which does not fit the real facts. Further, competitive price discrimination will generally be efficient and output-maximizing where feasible. None of this means competitive discriminatory pricing will always be feasible because in a competitive market individual firms would have a constant temptation to try to undercut the price to high-demand customers and grab a disproportionate share of them. But the existence of widespread price discrimination in competitive markets suggests that such price discrimination often is feasible. And the explanation probably lies in the fact that, in competitive markets, firms that just try to serve the high-demand consumers are driven out by reactive above-cost price cuts, as discussed next.

C. Why Competitive Price Discrimination Will Often Require Reactive Above-Cost Price Cuts

If market conditions completely bundle high- and low-demand purchases into one firm, or all high-demand purchasers are uninformed and price-insensitive, then a rival or entrant has no incentive to undercut a discriminatory price. But to any extent those conditions do not completely hold,¹⁷⁸ then a rival or entrant has some incentives to deviate from the output-maximizing price-discrimination schedule by offering high-demand buyers (such as adult moviegoers or hub customers) a lower price. That lower price can reap a disproportionate share of those customers and earn supracompetitive profits because the lower price would still exceed the sum of separate costs and a proportionate share of common costs. Professor Michael Levine, in his seminal analysis of price discrimination without

176. See Baumol, *supra* note 138 (manuscript at 21) (noting that this new analysis of price discrimination has forced him to recant his prior conclusion that such differences in route prices did indicate market power).

177. See Levine, *supra* note 138, at 6 (citing James D. Dana, Jr., *Advance-Purchase Discounts and Price Discrimination in Competitive Markets*, 106 J. POL. ECON. 395, 396 (1998)).

178. To any extent they do hold, however, they diminish competitive incentives for undercutting the output-maximizing discriminatory price. Partial fulfillment of these conditions thus reinforces the conclusion below that any increased short-run profits a firm might hope to earn by undercutting this discriminatory price are sufficiently small that they can be deterred by the fact that such cuts will provoke immediate or rapid price matching by rivals.

market power, assumed to the contrary that all firms, including airlines, that try to deviate from the optimal price schedule by offering lower prices to the high-demand buyers will “tend to disappear” because they will make less revenue.¹⁷⁹ And in his extension of Levine’s analysis, Professor Baumol at points tends to make the same sort of assumption that deviating from optimal price discrimination must sacrifice revenue.¹⁸⁰ But while the optimal price-discrimination schedule is revenue- and output-maximizing for the industry as a whole given the overall proportion of high- and low-demand buyers who exist, it does not follow that it is so for individual firms. If the incumbents stick to the optimal schedule, a deviating entrant can hope to profit by lowering prices to the high-demand buyers and attracting enough to make them a disproportionate share of its own customers. This is true even when the deviating entrant has to incur the same common costs because the optimal price schedule will reflect a price to the high-demand buyers that covers not only their separate costs but also a disproportionately high share of common costs. This, by definition, can be undercut by a lower price that remains above the sum of their separate costs and an equal share of common costs. Thus, if a slightly lower price to the high-demand customers shifts enough of them to the deviating firm that more of its customers are now of the high-demand variety, then at the new price its revenue will exceed its total costs, and it will enjoy an increase in profits.¹⁸¹

This does not mean that optimal price-discrimination schedules can never be maintained in competitive markets where high-demand buyers are not completely uninformed, price insensitive, or bundled with low-demand buyers. Rather, it means that the reason they are maintained is not that deviations from them are, standing alone, unprofitable. The reason is instead that, in a competitive market, incumbent firms would have to respond to any deviation by lowering their own prices to the high-demand

179. See Levine, *supra* note 138, at 14, 23-24.

180. See Baumol, *supra* note 138 (manuscript at 4-6). Elsewhere Professor Baumol recognizes that prices that exceed marginal costs will invite entry. See *id.* (manuscript at 18-20). But he does not link this observation to the point that a deviating entrant or rival can hope to get a disproportionate share of the high-demand buyers, a point that implies that deviation will be most attractive when getting such a disproportionate share is most feasible and when reactive above-cost price cuts will not deprive a deviating firm of its disproportionate share.

181. The situation is different in the case of slaughterhouses that charge more per pound for filet mignon than for lesser cuts of meat, on which Professor Levine focused much of his analysis. See Levine, *supra* note 138, at 14-16, 18. There, a deviating slaughterhouse that charges less for filet mignon cannot hope that such a price reduction to its high-demand buyers will increase the proportion of meat it produces that is filet mignon. Rather, it will still produce the same proportion of other cuts of meat but be unable to increase the price for those remaining cuts above prevailing market prices. Since it has lowered its price for filet mignon, but cannot increase its price for other cuts of meat, its total revenue will decline. And since the optimal price-discrimination schedule produced revenue that barely covered the common costs of slaughtering whole cows, this lower revenue will necessarily fail to cover the common costs, and such firms should indeed disappear.

buyers in order to retain a proportionate share of their patronage, for revenue from these high-demand buyers is necessary to cover common costs. This reaction is involuntary because the firms cannot stay in business unless they retain a proportionate share of high-demand buyers. Indeed, they will react in the same way to a decline in their proportion of high-demand buyers whether or not they noticed the deviation that caused it and even if they think their reaction will have no effect on the behavior of the deviating firm. Because these lower prices to high-demand buyers mean prices will now cover a smaller share of common costs, the incumbents will effectively face higher costs when deciding whether to incur the incremental common costs that are necessary to continue serving the low-demand buyers. Thus, in a competitive market, such deviations mean the incumbents will also have to raise prices to the low-demand buyers to continue to cover all common costs. This price increase does not mean that incumbents are exercising some market power over the low-demand buyers that they previously failed to exercise. Rather, it results because the costs of serving low-demand buyers have effectively increased. At the new price schedule, firms will not only incur common costs less often, reducing total output, but a smaller share of that output will be available for low-demand buyers. Without any market power, this reduction in output will raise prices for the low-demand buyers in order to balance supply and demand.

This change in the market's price-discrimination schedule means that the deviating rival or entrant will not enjoy, in the end, a disproportionate share of high-demand buyers. If that rival or entrant also serves the lower-demand buyers and is equally efficient, it will also have to raise prices to those buyers in order to cover common costs itself, and will thus go along with the general increase in market prices to the low-demand buyers. The deviating rival or entrant will thus end up (like the rest of the market) at a new price-discrimination schedule. But this new lower price discrimination will mean lower industry output and thus be inefficient and less profitable for everyone. Efforts to undercut the output-maximizing price-discrimination schedule will thus reap the deviating rival or entrant no additional profit unless other firms are slow to respond to the change in their proportion of high-demand buyers, and will lower its profits once they do respond. If deviation is rendered unprofitable by rapid rival response, firms or entrants will have little incentive to engage in it and thus the output-maximizing price-discrimination schedule can be sustained in competitive markets.

In short, the competitive practice that is necessary to maintain an output-maximizing price-discrimination schedule is precisely that incumbents will adopt rapid reactive above-cost price cuts that make it unprofitable for rivals or entrants to try to serve only high-paying customers. Rapidity is key, or deviating rivals or entrants can enjoy short-

term profits that will destabilize this price-discrimination schedule. This may explain why many retailers adopt announced policies of automatically (and retroactively) matching any lower price that its rivals may offer. Some have thought these policies are anticompetitive because they facilitate price discrimination or oligopolistic coordination.¹⁸² But retail markets seem far too unconcentrated and competitive to sustain any price discrimination or coordination that tries to raise prices above efficient cost levels.¹⁸³ A more likely explanation is that running a retail store involves incurring common costs every day to serve a range of high- to low-demand consumers, and the most efficient and output-maximizing method of covering those common costs is the widespread retail price discrimination we routinely observe, whereby consumers of high-end products pay a much higher markup than other consumers. But, to the extent high-demand retail buyers are not completely uninformed, price-insensitive, or bundled with low-demand buyers, such price discrimination would be subject to destabilization by retailers who undercut prices to the high-end customers unless reactive price cuts by incumbents are extremely rapid. Committing in advance to a retroactive price-matching policy is a way of assuring such rapidity without incurring the difficult task of monitoring numerous other retailers. In other industries where there are fewer rivals or entrants to monitor, or a longer lag time between deviating on price and actually obtaining customers, incumbents might not need such an automatic price-matching policy. Instead, they can rely on a general practice of adopting reactive above-cost price cuts when a deviating rival or entrant tries to undercut the prices to high-demand customers.

One might object that such price discrimination cannot be sustained on competitive markets given game theoretic considerations. The “game” is that each firm must make a choice between maintaining high prices to high-demand customers or deviating from them. The objection is that each firm would always choose deviating because this choice makes it better off no matter what choice other firms make. If it expects its rivals to maintain the high prices to high-demand customers, it is better off deviating because it gets a higher share of them and supracompetitive profits. If it expects its rivals to deviate by cutting prices to high-demand customers, it is better off doing the same because it has to in order to cover common costs.

182. See Edlin, *supra* note 166, at 529-31, 536-52 (arguing that such policies facilitate anticompetitive price discrimination); *id.* at 530 n.5, 531 n.9, 533 n.14 (collecting sources arguing that such policies can facilitate oligopolistic coordination).

183. See *supra* Section III.B.

TABLE 1. THE GAME THEORY OBJECTION

		Rivals	
		Maintain	Deviate
Firm	Maintain	0, 0	-2, 1
	Deviate	1, -2	-1, -1

Given this table of payoffs, for each individual firm, deviation seems a dominant strategy that makes maintaining the optimal price-discrimination schedule impossible.¹⁸⁴ This is true even though deviation by all firms makes each firm worse off by lowering output.

The problem with this objection is that it assumes there is some period where the deviating firm can charge lower prices to high-demand customers than its rivals. But if every firm has a policy of automatically and retroactively matching any price cut by its rivals, then no firm can ever offer a lower price than its rivals. The effect of an automatic price-matching policy is thus to eliminate the boxes where one firm has low prices and the other firms have high prices.

TABLE 2. PAYOFFS WITH IMMEDIATE PRICE MATCHING

		Rivals	
		Maintain	Deviate
Firm	Maintain	0, 0	N/A
	Deviate	N/A	-1, -1

The boxes where one firm deviates while its rivals do not thus never arise because deviating triggers immediate deviation by its rivals. Thus, the choice is between the boxes where both maintain and both deviate, and each firm will thus prefer to maintain the output-maximizing price-discrimination schedule.

With airlines, the price matching is not automatic and retroactive. But entry is foreseeable in advance, as is rival pricing, which needs to be announced weeks ahead of flight given the planning needs of customers. Thus, when the entrant is deciding whether to deviate by investing in entry that serves only nonstop customers, it knows in advance that, by the time it

184. See ERIC RASMUSEN, *GAMES AND INFORMATION: AN INTRODUCTION TO GAME THEORY* 19-22 (3d ed. 2001) (defining dominant strategy equilibria).

actually operates, incumbents will match its lower prices to those customers. Indeed, the entrant knows incumbents will be forced to do so in order to cover common costs. Likewise, when an existing firm announces lower prices to high-demand nonstop customers for a given flight, it knows that its rivals will have to respond by lowering their prices to those nonstop customers, and can do so well before flight time.¹⁸⁵ Indeed, they will respond immediately in order to retain enough nonstop customers to cover common costs whether or not they know what caused their drop in bookings from nonstop customers. So again, the two boxes where one firm can offer lower prices than its rivals drop out, and the choice is just between universal maintenance and universal deviation.

Even if the rival price matching is not immediate, the deviating firm will know that it will also take some time for its altered pricing to attract more high-demand buyers, and that it will enjoy at best a very short period of price advantage on the high-demand customers before its rivals are forced to match that deviating price. Rather than myopically considering only the positive consequences in that very short period, it will consider the consequences over the longer haul. If the additional profits during that very short period are outweighed by the losses of output, then deviating first will not improve the firm's fortunes. Instead, the payoffs will be as follows.

TABLE 3. PAYOFFS WITH DELAYED PRICE MATCHING

		Rivals	
		Maintain	Deviate
Firm	Maintain	0, 0	-1.1, -0.9
	Deviate	-0.9, -1.1	-1, -1

In this situation, maintaining the output-maximizing price-discrimination schedule will not be a dominant strategy, but it will be a Nash equilibrium. It will not be a dominant strategy because maintaining the price-discrimination schedule is not the best choice regardless of what the firm expects its rival to do.¹⁸⁶ If the firm expects the rival to maintain, then it is better off maintaining too. But if the firm expects the rival to deviate first, then it is better off deviating right away too. But it is a Nash equilibrium because "no player has incentives to deviate from his strategy given that the other players do not deviate."¹⁸⁷ That is, if the output-

185. In practice, hub-and-spoke airlines normally operate out of different hubs. Thus, rivals will probably have to enter that hub with more capacity in order to undercut nonstop prices.

186. RASMUSEN, *supra* note 184, at 19.

187. *Id.* at 33.

maximizing price-discrimination schedule is in place, no firm has incentives to deviate from it.¹⁸⁸

None of this is consistent with the assumptions of a perfectly competitive market, for that assumes (1) a “perfect divisibility of output” that elides the very common cost issue of concern and (2) that firms are all price takers who cannot lower prices because prices are at cost, and who cannot raise prices above cost without losing all their customers.¹⁸⁹ But such markets do not exist, and the above is consistent with the sort of workable competition that actually does exist in normal competitive markets.¹⁹⁰ Moreover, while firms that engage in competitive price discrimination are not price takers, their conduct is dictated by the market in the sense that their pricing behavior turns on the proportion of high-demand buyers available to them at each price and the size of their common costs, whether or not they noticed (or think they can affect) the rival pricing behavior that influenced that proportion.

One might relatedly object that the game theoretic argument described above for maintaining price discrimination is really an argument about coordination no different than an oligopoly model. But the crucial difference is that here the anticipated response by rivals is not strategic, but rather is forced by their need to cover costs and is thus involuntary. It does not depend on the reacting firm noticing the deviation that caused its proportion of high-demand customers to drop or on any expectation that its reaction will alter the behavior of the deviating firm or even be noticed by that firm. It is thus similar to the typical assumption in competitive markets that a firm cannot raise prices by constricting output because rivals will react by immediately expanding their output. In contrast, in an oligopoly that coordinates on supracompetitive prices, rivals have incentives not to follow a price cut immediately, because at the higher price they are earning positive economic profits and can continue to do so unless the deviating

188. This stylized table does not itself indicate, however, that mutual price discrimination is a unique Nash equilibrium because it suggests that, if all firms are deviating, no firm has an incentive to move first to the output-maximizing price-discrimination schedule because it cannot get a proportionate share of high-demand customers. If all firms are deviating, however, then demand from the high-demand buyers will exceed output and drive up prices. At that elevated price, competition to gain the greatest share of high-demand buyers will increase output until unused capacity is created that will give competitive firms incentives to price discriminate, which in turn will further increase output to the maximum level. See *supra* Section III.B.

189. See CARLTON & PERLOFF, *supra* note 19, at 57. Others also explicitly assume what seems implicit in most perfect competition models: that “producers have production functions that rule out increasing returns to scale.” VISCUSI ET AL., *supra* note 57, at 73. This assumes away common costs, which mean increasing returns to scale over certain incremental output ranges. Perfect competition models also assume perfect knowledge and continuous cost and demand curves, which again is inconsistent with the assumptions here.

190. See AREEDA & KAPLOW, *supra* note 132, at 35-38 (noting that perfect competition rarely exists, and discussing performance criteria for workable competition (like normal profits) and structural criteria (like many firms and low entry barriers), criteria which would be met in the markets under consideration).

firm can immediately expand capacity to take all their output. They will thus lower prices immediately only if they noticed the deviation that caused their decreased sales and think that their immediate reaction will send a strategic message to the deviating firm. Likewise, in a traditional oligopoly, firms must be able to observe rival prices and assess quality differences.¹⁹¹ In a market with competitive price discrimination, this is unnecessary. Firms need only observe their own proportions of high-demand and low-demand customers. If their proportion of high-demand customers has gone down, whether because of changed market conditions or rival deviation from the output-maximizing price schedule, their reaction will be the same. They will immediately lower prices to high-demand buyers to retain enough of them to cover common costs whether or not they know or care that a deviating firm caused this change in proportion or think they can send a message to the deviating firm. They will thus behave in the same way even when there are many more firms in the market. Since such firm behavior is not strategic in the sense of being chosen, designed to affect its rivals, or even aware of rival behavior, it seems more accurate to call the behavior competitive.¹⁹²

Two features make competitive price discrimination on hub-and-spoke flight systems different than in markets such as retailing consumer products or exhibiting movies. First, a retailer or theater need only maintain a proportionate share of sales to high-demand buyers. In contrast, demand and supply conditions may well mean that maintaining a hub-and-spoke system requires an airline to retain a disproportionate share of sales to customers who fly nonstop to or from their hub because their system focuses around a particular hub that rivals do not share. With this disproportionate hub share, incumbent hub-and-spoke airlines can more plausibly be accused of having market power that they are protecting with their reactive above-cost price cuts.

Second, airline hub-and-spoke systems of price discrimination are more vulnerable to being undercut by rivals who provide only the high-value slice of the system because an entrant can provide flights on a single route without incurring the common costs of servicing the rest of the hub-and-spoke system. Theaters or retailers may be somewhat less vulnerable to this

191. See *id.* at 254; George J. Stigler, *A Theory of Oligopoly*, 72 J. POL. ECON. 44 (1964).

192. Of course, we could define our terms to call this instead a mild form of "oligopolistic coordination," because firms are deterred from deviating by the anticipated (involuntary) response of their rivals, and then just conclude that in this case oligopolistic coordination turns out to be desirable, efficient, and output-maximizing. But in this context that definition would be both (1) misleading, since it conflicts with ordinary understandings that oligopolistic coordination is strategic and occurs among few firms, and (2) unhelpful, since it no longer would functionally correspond to an undesirable state of affairs. See *supra* note 137 (offering similar reasons to reject a definition of "market power" that would call this a desirable exercise of modest market power in a market with many rivals).

problem. A deviating theater may try to fill its seats entirely with adult moviegoers, but that will be difficult to do, both because they are often accompanied by children and because selling only to adults may not produce enough customers to fill all a theater's showtimes and cover common costs. Likewise, a deviating retailer may have difficulty getting the foot traffic to cover its common costs by selling only to high-demand consumers. Indeed, the same buyers may be high-demand consumers for some products (movie tickets for themselves or luxury products) and low-demand consumers for other products (tickets for their kids or commodity products) that they prefer to buy at the same time after making a trip to a single seller.¹⁹³ Further, the differences in firm prices for movie tickets or retail goods are sufficiently small that many high-demand buyers may not be willing to incur the relatively labor-intensive costs of becoming informed about them. Thus, while deviations will occur in theater and retail markets, they are less likely to be profitable, because it is harder to gain a disproportionate share of the high-end customers.

In contrast, an airline entrant can offer nonstop service on a single route without incurring the common costs of running a hub-and-spoke system at all. The differences in airline prices are also sufficiently large, and the search costs of discovering them (especially with the Internet) sufficiently low, that more high-demand buyers are likely to be aware of them. And even though any given passenger may take both high- and low-demand flights (if he sometimes flies as a business traveler and sometimes as a tourist), he normally does not need to book those flights with the same airline. True, airlines will try to reinforce brand loyalty with mechanisms like frequent flier programs in order to bundle those choices,¹⁹⁴ but it is unlikely this will be as successful as the bundling more naturally created by the costs of traveling to retailers. Thus, airlines are more likely to have incentives to deviate from the optimal price-discrimination schedule.

Accordingly, in a competitive market for hub-and-spoke systems, each hub-and-spoke airline will offer the price-discrimination schedule that maximizes the output of the entire hub-and-spoke system when they can, but will frequently have to deviate from that schedule by sharply reducing prices on nonstop flights to or from their hubs when less efficient entrants try to take just that slice of the market.¹⁹⁵ This above-cost price reduction

193. See *supra* Section III.B (explaining how in such cases sellers will compete to maximize buyer utility by offering discriminatory Ramsey pricing).

194. See *supra* Section III.B.

195. If an incumbent airline faces a more efficient entrant, it will not be able to lower prices enough to maintain a significant volume of hub passengers and will thus have to rely mainly on connecting passengers. This is how American Airlines dealt with the lower-cost competition provided by Southwest Airlines. *United States v. AMR Corp.*, 140 F. Supp. 2d 1141, 1181-82 (D. Kan. 2001). Such a great reliance on connecting passengers may well lower the efficiency of the

means these nonstop flights will cover a smaller share of the common costs of the system than before. When that is the case, entry that reduces prices for nonstop flights may not be desirable because (1) the price reduction does not stand alone but rather reflects a transfer payment to nonstop customers on that route from connecting customers and nonstop customers on other routes who benefit from the flights partly supported by demand from those connecting customers, and (2) the deviation from the optimal price schedule lowers total flight output across the hub-and-spoke system, including reducing the number of flight options for nonstop customers. Indeed, competition from such entrants that forces hub-and-spoke airlines to deviate from output-maximizing discriminatory prices makes all consumers worse off to the extent that they take both nonstop and connecting flights on different occasions.¹⁹⁶ In any event, firms in a fully competitive hub-and-spoke market would necessarily respond to single route entrants by lowering prices in order to continue to cover common costs. And if, as hub-and-spoke economics indicates, it is more efficient for one airline to provide both the nonstop and connecting flights,¹⁹⁷ then an entrant who provides only the nonstop flight simply cannot survive in a competitive market because a hub-and-spoke firm will be driven to undercut it with an above-cost price cut.¹⁹⁸ Those price cuts will then drive such entrants out and allow the incumbents to raise prices on that route back to the price that matches the output-maximizing price schedule for the hub-and-spoke system.

In short, the observed pattern of single route entry, reactive above-cost price cuts by hub incumbents, exit by the single route entrant, and restoration of higher prices can be explained by fully desirable, competitive behavior. One need not assume that the incumbent airline must have monopoly power that it is trying to protect through strategic pricing. Indeed, the nonmonopoly explanation seems more consistent with the empirical evidence that the airline industry has not only failed to enjoy monopoly profits, but has been unable to sustain even a competitive rate of return for any five-year period since deregulation.¹⁹⁹ The competitive explanation also helps explain why firms do not cut prices any more when a

hub-and-spoke system as a whole, but it does shift hub passengers to a clearly more efficient provider.

196. See Bliss, *supra* note 163, at 387-88 (proving the parallel point that, when grouping products in one retail store is efficient, competition from specialty stores causes deviations from optimal pricing that make consumers worse off).

197. See *supra* Section III.A.

198. See Bliss, *supra* note 163, at 387-88 (proving the parallel point that if the costs of traveling to a store make it efficient to buy high- and low-demand products at the same retail location, and costs are convex and homogeneous, then a specialty store that sells only the high-demand goods cannot survive in a competitive equilibrium).

199. See Dorman & Baumol, *supra* note 147, at 4; see also Baumol, *supra* note 138 (manuscript at 8) (noting that airline investments earn a lower return than stock indexes).

newly formed entrant enters a route than they do when an existing airline does.²⁰⁰

This analysis undermines the intuition that something nefarious is going on when an incumbent airline lowers nonstop prices to above-cost levels in response to single route entry and then raises them again after entry. Both price changes can be explained by the simple reality that the fact of entry (at a particular price) has changed the market price on the nonstop route and thus requires a readjustment of the price-discrimination schedule, and the fact of exit makes the old price-discrimination schedule optimal again. This analysis turns on its head the intuition that reactive above-cost price cuts are used to foil the predictions of contestable market theory. To the contrary, if airline markets are in fact contestable, then such contestability will *force* airlines to engage in revenue-maximizing discriminatory pricing to cover common costs,²⁰¹ which can only be maintained through reactive above-cost price cuts.

Perhaps a more profound implication is that this analysis means that—even if proponents were entirely correct about the predicted effects of a restriction on reactive price cuts on individual airline routes—those predicted effects would likely be undesirable. If proponents are correct, their restrictions will lead to lower everyday prices on nonstop flights from concentrated hubs. But obtaining these benefits means deviating from the price-discrimination schedule that maximizes the overall output of the hub-and-spoke system. There is no particular reason to think that overall result would be desirable. It would be similar to legislating lower prices for adult movie tickets than the unregulated market would produce. Adult moviegoers who buy at those lower prices will be better off, but prices would have to rise for nonadults. Further, the overall output of movie exhibitions would go down since the new price schedule would no longer be the one that optimizes output. Likewise, even if the restrictions lowered prices for nonstop hub flights, that would make nonstop fliers from hubs to spokes better off, but raise prices on the rest of the hub-and-spoke system. There is no particular warrant in antitrust law for imposing such a distributional transfer by legally restricting competitive above-cost pricing. Even left standing alone, these distributional effects are likely to involve an undesirable shift from low- to high-income customers. Worse, this distributional effect will have been purchased at the cost of imposing a reduction in the overall output of flights between cities connected by hub-and-spoke systems. That will lower total social efficiency and aggregate

200. See Dennis W. Carlton & Gustavo E. Bamberger, Docket No. OST-1998-3713-1709, Reply Comments to the Department of Transportation's Proposed Enforcement Policy Regarding Unfair Exclusionary Conduct in the Air Transportation Industry 5 (Sept. 24, 1998), at http://dmses.dot.gov/docimages/pdf30/43821_web.pdf.

201. See Baumol, *supra* note 138 (manuscript at 1-3).

consumer welfare. Indeed, this will make *all* customers worse off to the extent that they take both nonstop and connecting flights at different times.

IV. RESTRICTING ABOVE-COST PRICE CUTS HAS ADVERSE EFFECTS EVEN WHEN THE INCUMBENT DOES HAVE MARKET POWER AND IMPLEMENTATION DIFFICULTIES ARE IGNORED

Even if we assume incumbent market power has been independently established without relying on evidence of discriminatory pricing or reactive price cuts, restrictions on reactive above-cost price cuts still will normally be undesirable. The traditional argument for thinking so stresses their administrative difficulties, which are formidable and even underestimated. But let me defer those issues until Part V. Here, I am interested in addressing the issue whether, even if we assume away any implementation difficulties, the proposed restrictions on reactive above-cost price cuts would be desirable.

Although the proposed restrictions on reactive above-cost price cuts differ in their details, they all would effectively set a floor on incumbent pricing after entry.²⁰² Professor Edlin would set that floor at the incumbent's pre-entry price.²⁰³ Professor Williamson would instead ban an incumbent from expanding output after entry.²⁰⁴ But every output ceiling implies an associated price floor. The Williamson rule would allow the incumbent to cut prices, but only to the extent necessary to maintain output after the entrant has added its own output to the market. Thus, although the Williamson output ceiling would allow prices lower than the Edlin approach, it does set an effective floor on post-entry incumbent pricing. The European doctrine in *Compagnie Maritime* and the proposals of the U.S. Departments of Transportation and Justice would effectively set a price floor at the level that "clearly" or "substantially" (or in the EU doctrine maybe "selectively") falls below the price that would maximize the incumbent's short-term profits after entry.²⁰⁵ Likewise, Professors Ordober

202. Professor Baumol's proposal, which would not restrict reactive above-cost price cuts but would require that they be quasi-permanent, is analyzed separately below in Part VI.

203. Edlin, *supra* note 6, at 945-46.

204. See Williamson, *supra* note 7, at 295-96, 333-36.

205. See *supra* Section I.A. The Departments would also set a ceiling on output expansions that fail this short-term profit-maximization test, but this output ceiling also implies an associated price floor. Which Department's proposal sets the higher price floor may depend on the circumstances. For example, if it were 100% certain that the incumbent could make a 1% higher profit with a higher price, then the Department of Justice position would require at least that price (since it would "clearly" increase profits), but the Department of Transportation position would not (since it would not increase profits "substantially"). Alternatively, if it were 51% certain that an incumbent could make a 50% higher profit with a higher price, then the Department of Transportation position would require at least that price, but the Department of Justice position would not. In general, however, one would expect the proposals largely to track each other since

and Willig and others had earlier proposed a similar test without the “clearly” or “substantially” or “selectively” qualifier.²⁰⁶ Since an entrant adds output to the market, the incumbent’s normal short-term profit-maximizing response to entry would be to constrict output somewhat, thus usually indicating a somewhat higher price floor than the Williamson rule, which allows the incumbent to lower prices further to maintain its pre-entry output.²⁰⁷ In any event, whichever sets the lowest price floor, all of them effectively set some floor on post-entry incumbent prices.

Such a restriction on above-cost price cuts by definition cannot protect an entrant who is not less efficient than the incumbent.²⁰⁸ Nor do the proponents claim their restrictions would protect entrants who are just as efficient as, or more efficient than, the incumbent. Rather, they focus on the claim that protecting less efficient entrants is desirable. Their essential claim is that the restrictions will either encourage additional entry by these less efficient entrants, or prompt incumbents to lower pre-entry prices (or expand pre-entry output) to avoid such entry, either of which will enhance consumer welfare and allocative efficiency by lowering prices below their normal monopoly levels.²⁰⁹ At points, some proponents also suggest that, while these encouraged entrants may initially be less efficient, if protected by the proposed restrictions, they may be able to stay in the market long enough to become just as efficient as the incumbent.²¹⁰ Assessing these claims thus requires comprehensively assessing the effects of the proposed restrictions on the likelihood and consequences of each type of possible entrant, and on the behavior and creation of incumbents.

Section A begins with the proponent’s paradigmatic case: entrants who are (and will remain) less efficient than the incumbent. It notes a point that proponents have neglected: Some of these less efficient firms would have entered with or without the restrictions. For them, the effects of the restrictions would be entirely adverse. The restrictions would raise post-

the more substantial the expected profit difference, the more likely it is to be clear that some profit is being sacrificed.

206. See Ordovery & Willig, *supra* note 71, at 9-10, 15-16; see also *supra* note 71 (collecting other prior authorities proposing a similar test, even when prices are above cost).

207. If demand increased sufficiently, the profit-maximizing response could be increasing output, which might suggest the Williamson rule would require a higher price floor to prevent pre-entry incumbent output from rising. But to avoid this result, Williamson ultimately makes his test one of “demand-adjusted” output. See *infra* Part V (discussing other complications this raises). In theory, the U.S. Departments’ approach might impose a lower price floor because they only ban prices that are “clearly” or “substantially” below the profit-maximizing level. But maintaining output in the face of an entrant’s addition to market output will normally more than satisfy this test. Moreover, Williamson also includes his own version of a clearly-or-substantially qualifier by allowing a ten percent increase in output over the demand-adjusted prediction in the hopes that this will circumvent problems with ascertaining demand-adjusted output. See *infra* Part V.

208. See *supra* Part II (defining costs to satisfy this condition).

209. See Edlin, *supra* note 6, at 945-49, 973-78; Williamson, *supra* note 7, at 308.

210. See Edlin, *supra* note 6, at 975 & n.95, 977; Williamson, *supra* note 7, at 296, 298 n.43, 303-04, 313.

entry prices, thus lowering output, harming consumer welfare, and causing allocative inefficiency. Further, the restrictions would cause a shift of production to less efficient firms, a loss of incumbent efficiency, and a wasteful infliction of uncompensated transition costs. Other less efficient firms might be encouraged to enter by the protection the restrictions offer. But these less efficient entrants will be inevitably driven out when—by passage of time or loss of monopoly power—any restriction on reactive price cuts by the more efficient incumbent expires. Since long-run returns are impossible, the only encouragement would be that these restrictions can increase the length of the short-run period when they remain in the market. But if the capital costs of entry are high, they cannot be recouped with such short-run returns. And if the capital costs of entry are low, then less efficient entrants would often enter anyway. At best, the restrictions may provide some weak encouragement to less efficient entrants when entry costs are in an intermediate range so that the additional profits from prolonging the short-run period provide the marginal increment necessary to make total short-run entrant profits cover entrant costs. Further, while such entry may well lower prices from pre-entry levels for those consumers who buy from the entrant in the short run, it can also give incumbents perverse incentives to *raise* post-entry prices to speed the day when the restriction expires, which would raise prices for the majority of consumers. Thus, even when the restrictions do encourage additional less efficient entry, the net effects on consumer welfare and allocative efficiency will be mixed. Any encouraged entry would also shift production to less efficient firms, wastefully impose uncompensated transition costs, and lower the efficiency of incumbents.

In short, even if one focuses only on less efficient entrants, the overall effects of the restrictions are almost certainly negative. Where the less efficient entrant would have entered anyway, there will be negative effects on consumer welfare and productive efficiency. Where the restriction encourages the less efficient entrant to enter, there will be a mixed short-term effect on consumer welfare and negative effects on productive efficiency.

But in fact one cannot assess the full effects of the proposed restrictions by limiting one's consideration to less efficient entrants. Rather, as Section B points out, one must also consider the effects (ignored by proponents of these restrictions) on entrants who are just as efficient as, or more efficient than, the incumbent. For such entrants, the effects of the restrictions are unambiguously adverse. They raise post-entry prices, lower output, harm consumer welfare, and lessen allocative efficiency. Further, the restrictions make the mix of entrants less efficient by increasing the returns to inefficient entry and by lessening the returns to successfully creating an entrant who is more efficient than the incumbent.

Section C considers the possibility that entrants will become more efficient than the incumbent over time. It concludes that this is often undesirable because it frequently depends on a decrease in the incumbent's efficiency. If, in contrast, it is achieved solely by an increase in entrant efficiency, then the restrictions should be unnecessary because capital markets would fund such entrants anyway. Further, the post-entry effects of the restrictions for such entrants are entirely adverse.

Section D addresses the effects of the proposed restrictions on pre-entry incumbent behavior. It concludes that it is doubtful the restrictions will induce incumbents to lower pre-entry prices, and that even if they do so, such a regime of enforced limit pricing is legally inconsistent with the argument for banning reactive above-cost price cuts. More important, proponents have neglected to take into account that, by lowering the rewards for creating an incumbent that is more efficient than other market options, the restrictions reduce the incentives for the innovation and investment necessary to create those more efficient incumbents in the first place.

Section E summarizes the effects and concludes that the trade-offs almost certainly cut against the proposed restrictions even if one ignores implementation difficulties. And Section F rebuts the possibility that the problems with the restrictions can be avoided by modifying the market-power requirement.

A. *Effects on Likelihood and Consequences of Less Efficient Entry*

I begin by considering the effects of the restrictions on the likelihood and consequences of entry by firms that are less efficient than the incumbent throughout the period of any restriction on reactive above-cost price cuts. Such less efficient entrants form the centerpiece of the proponent's arguments for restrictions.²¹¹ An entrant can be less efficient because its costs are higher than the incumbent's, because its quality is lower at the same cost, or because it offers a cost-quality trade-off that consumers find less attractive than the incumbent's. Since the last two amount to saying the entrant has higher costs in delivering the level of quality that consumers prefer, I will call all three the case of a higher-cost entrant. In the long run, the incumbent firm with a cost advantage can drive such entrants out of the market by cutting its prices to a level above the incumbent's costs but below the entrant's costs, which the entrant cannot profitably match. Likewise, an incumbent with a quality advantage can

211. See Edlin, *supra* note 6, at 944, 955-60, 962-63, 965, 973-78; *supra* text accompanying notes 1-14.

drive the entrant out of the market in the long run by pricing at its own cost, giving consumers either higher quality at the same cost or a quality-cost trade-off they prefer.

Some of these less efficient entrants would have entered even without the restrictions because the short-run profits of doing so are sufficiently enticing. In those cases, the restrictions are unambiguously adverse for both consumer welfare and productive efficiency. Other less efficient entrants might have been induced to enter by the restrictions. In their case, the restrictions will have mixed consequences for consumer welfare but a negative effect on productive efficiency. Further, the restrictions will encourage additional entry by relatively few less efficient entrants because the restrictions will eventually expire and thus cannot protect less efficient entrants in the long run.

1. *Consequences for Less Efficient Entrants Who Would Have Entered Without Any Restriction*

Many less efficient entrants would have entered even without the protection of a rule that restricts above-cost price cuts. For cases involving such entrants, the consequences of the restriction will be unambiguously negative.

a. *Why Less Efficient Entrants Often Enter Without Any Restriction on Reactive Above-Cost Price Cuts*

Although entrants who are just as efficient as, or more efficient than, the incumbent will not be deterred under a cost-based test, the converse does not follow that all less efficient entrants will be deterred. To the contrary, less efficient firms will often enter a monopoly market under a cost-based test even without the protection of a restriction on above-cost price cuts. After all, by hypothesis, the preexisting market was priced at supracompetitive levels. Thus, even a less efficient entrant can offer a lower price that exceeds its costs and reap supracompetitive profits in the short run.

True, in the long run, the more efficient incumbent will be able to drive out the less efficient entrant with above-cost price cuts. But the short run may not be so short. The longer it lasts, the greater the entrant's profits will be. And the longer lasting any price cut must be to drive out an entrant, the more likely the incumbent would find it more profitable to accommodate entry at higher prices rather than trying to cut prices to drive out the entrant. This can be obscured if the airline industry is the paradigmatic case one has in mind. While the airline industry proves a poor paradigmatic case because

reactive price cuts there probably do not protect market power at all,²¹² the airline industry does have a combination of features that make it more susceptible to driving out less efficient entrants with very short-term price cuts. Namely, in the airline industry, incumbent capacity is easy to expand, and buyers cannot realistically engage in significant long-term contracting or storage. In markets lacking this combination of features, a reactive price cut designed to drive out entrants cannot be nearly so temporary, for the following reasons.

If capacity cannot easily be expanded, then it may take the incumbent a significant period to expand output enough to drive out a less efficient entrant.²¹³ True, for physical products made in plants, the incumbent may maintain some excess capacity for just this purpose. But the costs of doing so may not be worth bearing.²¹⁴ Moreover, even in such a plant, expanding capacity may not be as easy as turning on a switch. Extra personnel have to be added or trained, or if the incumbent has also kept excess workers idle, their skills will be rusty. These problems are likely to be even greater in service industries. The airline industry is unusual in this regard because the relevant capital goods and personnel are so easy to move to a targeted market.

Even if the incumbent can rapidly expand output, buyers will have incentives to respond to any price cut they anticipate is temporary by stockpiling as much as possible of the good. Thus, rather than the incumbent's expanded output replacing purchases from the entrant, buyers have incentives to buy as much as they can from both and stockpile their purchases. This effectively makes any temporary price cut more permanent. This is not a feasible consumer reaction in the airline industry because future travel needs are sufficiently uncertain that it is hard to stockpile too many tickets. But it seems far more likely to be a feasible reaction in markets where the incumbent is just turning on plant capacity to make a physical good, which was the one case where incumbent-output expansion seemed likely to be faster than entrant-output expansion.

212. See *supra* Part III.

213. This generally is not an issue when, instead of protecting market power, the incumbent is reacting to an entrant who is undercutting competitive price discrimination because in that case the incumbent does not need to expand output. It just needs to reallocate output now going to low-demand buyers. Indeed, overall output will likely decline. See *supra* Part III.

214. Williamson assumes that under any rule the incumbent will invest to maintain enough excess capacity to be able to reduce entrant profits to zero. See Williamson, *supra* note 7, at 294, 297-98, 310 n.66, 314. But in many markets, this may be too costly to be profitable at all, and in all markets it involves a trade-off between pre-entry profits and post-entry hazards that may not be worth making. Williamson's contrary conclusion is based on what he admits is the "arbitrary assumption" that incumbents strictly prefer avoiding post-entry hazards to earning pre-entry profits. *Id.* at 314. There is no reason to think this assumption is accurate, and thus incumbents often will not have sufficient excess capacity on hand.

Finally, in any market where buyers engage in long-term contracting, an entrant facing the prospect of a reactive price cut can try to contract with enough buyers to assure its survival for long enough to recoup the costs of entry. This is not so feasible in the airline industry, where most purchasing is done on an effective spot market for each trip.²¹⁵ But it is feasible in many markets. Williamson recognizes long-term entrant contracting is possible, but assumes it will be rare for three reasons. First, he assumes long-term contracting is generally inefficient. But in many markets it is used, suggesting it is efficient in those markets. Second, he assumes customers will not want to commit themselves unless the entrant has committed itself by incurring fixed costs. But any long-term contract can be made contingent on the entrant incurring those costs or initiating actual entry. Third, he assumes the dominant firm will contest these pre-entry sales. True, but if so, then the “temporary” price cut will be even less temporary, extending to pre-entry periods and beyond if the incumbent itself offers long-term contracts to compete. At the extreme, the dominant firm will have to keep offering competitive prices all the time to fend off entrants.

Limits on these factors do, however, mean that sometimes relatively short-term price cuts can drive out less efficient entrants. Stockpiling may be impossible or costly if storage expenses are high, goods are perishable, services are time-sensitive, or future needs are difficult to estimate. Stockpiling will also be limited if buyers mistakenly expect the price cut to be permanent. The more difficult or costly storage is, and the more mistaken consumer expectations are, the more any market resembles that of an effectively nonstorable good like airline flights.

Long-term entrant contracting will also be limited to the extent it has inefficiencies or buyers face collective action problems. Markets with one buyer face no collective action problem because that single buyer can itself determine whether the entrant stays in the market. Thus, a single buyer would compare the entrant’s long-term contract price to the expected incumbent price stream, which features a temporary cut and then monopoly prices. But markets with many buyers face a collective action problem because each individual buyer will correctly figure that its single long-term contract will not significantly affect the odds that entry will occur or be successful. Judge Frank Easterbrook concludes that any collective action problem can be avoided by having each buyer enter a long-term contract with the entrant at a price below pre-entry prices and contingent on the

215. Even in the airline industry, though, corporations can and do negotiate for long-term discounts from regular prices. The main problem in that industry has been that the incumbent airlines are the ones with those contracts, thus making it harder for entrants to break in. *See United States v. AMR Corp.*, 140 F. Supp. 2d 1141, 1180 (D. Kan. 2001).

entrant getting enough commitments to be successful.²¹⁶ Alas, this does not avoid the collective action problem. An individual buyer's decision to join such a contract cannot make the buyer better off unless it meaningfully changes the odds of successful entry, and this is true no matter what the individual buyer hypothesizes the end result will be. If the entrant ultimately will not enter, joining such a contract gains the buyer nothing. If the entrant will enter but be driven out, then the entrant will not supply the product in the long run, and, in the short run, the buyer will be better off accepting the incumbent's temporary price cut to a level below the entrant price. If the entrant will enter and succeed, the buyer need not join the contract to get the benefit of entrant prices in the long run, and, in the short run, the buyer will still be better off accepting the incumbent's temporary price cut.²¹⁷ Thus, although buyers collectively have an incentive to enter long-term contracts with entrants to encourage their entry, buyers individually may not have such an incentive in markets with many buyers.²¹⁸ The greater the buyers' collective action problems and the shorter the term of an efficient contract in their market, the more other markets will resemble markets with little long-term contracting, like the airline industry.

Accordingly, it is hardly the case that less efficient entrants who could undercut a monopoly price would always enter the monopoly market regardless of the prospect of reactive above-cost price cuts. It is simply the case that many of them would. And in these cases, the effects of the proposed restrictions are unambiguously undesirable, as shown next.

216. See Easterbrook, *supra* note 17, at 270-71; see also CARLTON & PERLOFF, *supra* note 19, at 336-37 (assuming also that buyers would be willing to contract with the entrant at a price below pre-entry prices).

217. This is the difference between the situation here and the typical situation where collective agreements are successful. Here, while buyers are better off with a successful collective agreement than without one, they are even better off if the collective agreement occurs without their involvement. Ironically, a nonnegotiable agreement that required *unanimity* would be more likely to be adopted because then joining would be costless in the sense that buyers could not hope to do better outside the agreement than in it. But unanimity will be hard to achieve in markets with many buyers, and in practice such agreements cannot be truly nonnegotiable, which means any unanimity requirement creates holdout problems. Namely, the last buyer has incentives to demand that, in exchange for joining, it get preferential terms that amount to expropriating a greater share of the gains of the successful collective agreement. And this prospect will give all firms an incentive to put off agreeing so that they can be the last firm, thus recreating the collective action problem.

218. Even if there is a multitude of consumers, there may be sufficiently few buyers up the distribution chain—like retailers or wholesalers—to enable them to enter into long-term contracts with entrants. Easterbrook, *supra* note 17, at 271. On the other hand, retailers or wholesalers also have incentives to enter into Coasian bargains with the monopolist to split the supracompetitive surplus rather than eliminate it because increased costs can be passed on to consumers in higher prices, and the resulting decreased volume can be made up for by getting a share of the monopoly profits. See 4 PHILLIP E. AREEDA ET AL., *ANTITRUST LAW* ¶ 943b, at 204-06 & n.4 (rev. ed. 1998).

b. *The Undesirable Consequences*

For those less efficient entrants who would enter even without a restriction on post-entry incumbent prices or output, the restrictions can have no positive effect on their likelihood of entry. Rather, the only consequences will be on post-entry price competition, and those will be unambiguously negative.

The restrictions will all limit the post-entry competition that otherwise would have occurred between incumbents and less efficient entrants. Below-cost price cuts would be prohibited even without the restrictions. Thus, where the restrictions have bite, they will prevent incumbents from making above-cost price cuts that lower their price as much as they otherwise would have.²¹⁹ Indeed, an unrecognized cost of the restrictions is that they would give incumbents affirmative incentives to *raise* prices. The reason is that these restrictions would all expire once the incumbent loses enough market share to fall below whatever threshold is deemed necessary to establish monopoly or market power in that market.²²⁰ Accordingly, since the incumbent can drive the less efficient entrant out after the restriction expires but not before, it has perverse incentives to *lose* market share to the entrant as rapidly as possible to bring closer the day when the restriction expires and it can drive the entrant out and restore monopoly pricing. One natural way to lose market share will be to increase prices to a level that is more profitable on any sales that the incumbent does make. The incumbent will even have incentives to raise prices *above* its short-term profit-maximizing level because that speeds the return of long-run monopoly profits. Note the irony. The concern prompting restrictions on above-cost price cuts is that the incumbent might lower prices in ways that sacrifice short-run profits in order to reap long-term profits from excluding the entrant. But such restrictions can instead cause the incumbent to *raise* prices in ways that sacrifice short-run profits in order to reap long-term profits from excluding the entrant. This perverse incentive will exacerbate

219. Depending on market circumstances, it might be that the price floors set by the Williamson or short-term profit-maximization rules are below the price an unrestricted incumbent would want to charge post-entry anyway. In those cases, though, the restrictions have no bite.

220. Edlin stipulates that his proposed price floor applies only "until the entrant's share grows enough so that the monopoly loses its dominance." Edlin, *supra* note 6, at 945, 968-69. Williamson applies his rule only to dominant firms, which he defines as having a market share of at least sixty percent and enjoying significant entry barriers. See Williamson, *supra* note 7, at 292-93. Although Williamson's initial statement of his rule also applied to collusive oligopolies, *see id.*, he later recognized that applying his rule to such cases would have the undesirable effect of aiding oligopolistic coordination and thus seemed to abandon that extension, *see* Williamson, *supra* note 80, at 1195. Likewise, U.S. and European antitrust laws and the proposed Department of Transportation regulation all require some level of monopoly or dominant market power. *See supra* Section I.A.; *infra* Section IV.E.

the tendency of the proposed above-cost floors on incumbent post-entry prices to raise prices and harm consumer welfare and allocative efficiency.

Consumers buying from the entrant will also pay higher prices than they would have paid without the above-cost floors on incumbent post-entry pricing. This is because, with the protection of the incumbent price floor, the entrant has little incentive to lower prices all the way down to its costs. Under the Edlin rule, the entrant will not offer any price below a twenty percent discount from pre-entry prices since it knows the incumbent cannot cut prices. Under the Williamson or profit-maximization price floors, the entrant has incentives to charge a price just below that price floor, even though unrestricted competition would have driven it to price lower. Even if the entrant is not initially sure just where the incumbent's price floor will be, the entrant can reveal that floor by setting its opening price high, and then very slightly undercutting each incumbent price cut until it arrives at a price just below the lowest price the incumbent can charge. Thus, not only will the entrant's ultimate price be no lower than a price just below the price floor, but the restrictions will give the entrant incentives to set its initial prices even higher to reveal that price floor.

In short, those who purchase either from the incumbent or the less efficient entrant will pay higher prices. This harms consumer welfare. It also harms allocative efficiency since the precluded lower prices would have been above cost.

The effects on productive efficiency are also unambiguously adverse. Where they have bite, the restrictions will prevent the more efficient incumbent from expanding its output as much as it otherwise would have, thus shifting production to the less efficient entrant. This shift of post-entry output to a less efficient producer alone necessarily lowers productive efficiency.

Further, unless market demand sharply increases with entry, the incumbent will have to lower its output significantly from pre-entry levels because the entrant is taking a large share of market output and the restrictions generally prevent the incumbent from lowering prices in order to maintain its old output. This is certainly true under the Edlin rule, which forbids any reduction in pre-entry prices. It also follows under a short-term profit-maximization rule even if we assume that, both before and after entry, the incumbent monopolist sets a short-term profit-maximizing price that implies subcompetitive output levels. The reason is that whatever output the entrant takes away causes a leftward shift in the incumbent's residual demand and thus (absent an offsetting increase in total market demand) implies that a lower incumbent output will maximize its short-run profits. An even more dramatic reduction in output will result if we take into account two additional factors. First, sometimes the pre-entry price will be an (unsuccessful) limit price rather than a profit-maximizing price, and

thus the restriction can affirmatively require the incumbent to *raise* prices to comply with the post-entry price floor. Second, the restriction, as noted above, gives incumbents perverse incentives to charge a post-entry price *above* the profit-maximizing level to speed the end of the restriction.

Although the Williamson rule does not require a post-entry output reduction, it will often induce one. After all, it sets a ceiling on output, so output can only stay the same or go down. On average, then, incumbent output has to decline somewhat. More important, the incumbent has affirmative incentives to reduce output in any case where the output ceiling actually protects an entrant from being driven out of the market in the short run.²²¹ There are three reasons for this. First, reducing output will likely increase the incumbent's short-term profits given that the entrant is now taking up some market demand. Second, where maintaining output cannot drive out this entrant and restore monopoly profits, the incumbent has no reason to sacrifice short-term profits by maintaining output. Third, to the contrary, it is *reducing* output that will bring closer the day when the incumbent's market share erodes sufficiently to lift the restriction and allow the incumbent to drive out the entrant. This means that under the Williamson rule, the incumbent who is prevented by the output ceiling from driving out an entrant actually has incentives to speed the day when the rule expires by pricing *above* the short-term maximizing price, which means setting output below that level. The result is that, in any case where it actually has bite—that is, actually protects the entrant from being driven out by above-cost prices—the incumbent will set the same short-term price under the Williamson output ceiling as under a profit-maximizing price floor. Williamson sees the first factor but apparently not the other two and, in any event, effectively excludes all of them from his model by simply assuming that in response to entry the incumbent will always set the maximum output allowed by the legal rule.²²² But we must assume incumbents will be dynamic not just in their responses to entry but also in their responses to legal rules that frustrate efforts to make entry unprofitable. Thus, the Williamson rule will produce on average a reduction in post-entry output and, in fact, will do so in every case where the rule prevents the incumbent from driving out the entrant.

To the extent the restrictions do make the incumbent reduce its output from pre-entry levels, this subjects the incumbent to a wasteful process of contracting production during the restriction period, which it then has to turn around and expand after the restriction expires. That may entail costly and disruptive layoffs, contractual breaches or changes, idling and

221. Williamson himself assumes his rule would never protect a less efficient entrant, but he is mistaken for reasons explained *infra* text accompanying notes 231-234.

222. See Williamson, *supra* note 7, at 294-95 & n.35, 297-98, 310 n.66.

maintaining capacity, building renovations and the like. Such contractions and closings are a necessary cost of competitive markets, where they have the virtue of signaling when resources should switch from one firm or industry to another. But they constitute sheer waste when a more efficient firm is being forced by regulation to mothball capacity that ultimately will return to the market. Even when the infliction of these transition costs does not affect the operating efficiency of the incumbent, they nonetheless reflect real costs that will be visited on owners, workers, and others who contract with the incumbent. This will increase the costs of contracting with the incumbent, thus elevating the contract prices the incumbent must pay and reducing the returns for having created an efficient incumbent.

Indeed, if the restriction causes a post-entry reduction in incumbent output, this will probably affirmatively reduce the incumbent's operating efficiency for various reasons. First, to the extent the incumbent's efficiency advantage results because of economies of scale or scope that still apply at large outputs, a reduction in its scale or scope will make it less efficient.²²³ Second, the incumbent has presumably selected a plant size that minimizes the short-run costs of producing its pre-entry output. Thus, any decline in output increases its short-run costs.²²⁴ Third, because the restriction on reactive price cuts may require the incumbent to mothball capacity and layoff workers in the short run, it may disrupt an efficient operation. Machines that were well-oiled may become rusted, or new workers may need to be hired and trained. When full production starts up again, the costs may thus be higher or the quality lower. If any of these three factors hold, then, a restriction that causes the incumbent's output to drop will also decrease its productive efficiency. That would mean that the restriction would effectively have shifted the *entire* market to less efficient production: either to the less efficient entrant or to an incumbent who is less efficient than it otherwise would have been.

In sum, for those less efficient entrants who would have entered without any post-entry above-cost floor on incumbent prices, all the restrictions would inflict harm to consumer welfare, a loss of allocative efficiency, a loss of productive efficiency, and the wasteful imposition of uncompensated transition costs.

223. An economy of scale results when average costs for a product fall as firm output increases, whereas an economy of scope results when two products can be produced more efficiently together than separately. See CARLTON & PERLOFF, *supra* note 19, at 35-40, 50-52.

224. See Williamson, *supra* note 7, at 297, 300-02, 309-10 (assuming that the incumbent plant size minimizes the short-run costs of making the pre-entry output, so that any decrease or increase in incumbent output necessarily reduces its efficiency and raises its costs).

2. *Effects for Less Efficient Entrants Whom the Restrictions Encourage To Enter*

Other less efficient entrants might be encouraged to enter because the restrictions set an above-cost floor on incumbent post-entry pricing. But this encouragement will be relatively weak because such restrictions cannot prevent such less efficient entrants from being driven out of the market in the long run. Where the restrictions do encourage entry by less efficient entrants, the consequences will be mixed. Consumers who buy from the entrant will pay lower prices than they otherwise would have. But the majority of consumers stuck buying from the incumbent may pay more. Further, productive efficiency will suffer and wasteful uncompensated transition costs will be imposed.

a. *Why Restrictions on Reactive Above-Cost Price Cuts Can Provide Weak Encouragement to Entry by Less Efficient Firms*

Less efficient entrants will sometimes be encouraged to enter by restrictions that set an above-cost floor on incumbent post-entry pricing. The reason is that such restrictions can effectively lengthen the short-run period when a less efficient entrant can hope to sell at prices that exceed its own costs. This will sometimes provide the marginal increment of additional profits that the less efficient entrant needs to make its total expected short-run profits higher than the capital costs of entry.

But this encouragement will be weak because the additional increment is relatively small and short-term. As noted above, less efficient entrants will often be able to survive in the market for *some* short-run period. The restrictions will increase the prospective profits from entry by increasing the length of this short-run period. But the restrictions cannot offer less efficient entrants any long-term protection. The Edlin and Williamson rules would expire in twelve to eighteen months. At that point, the more efficient incumbent can offer above-cost price cuts that will drive the entrant out. Further, all the restrictions would expire once the entrant expands enough to deprive the incumbent of whatever market share is necessary to establish its monopoly or market power. That may be far less than twelve to eighteen months when the incumbent cannot match an entrant price that undercuts it. Consumers are likely to switch rapidly to the lower-priced entrant, especially when (as under the Edlin rule) the entrant price is a full twenty percent below the incumbent price floor. The incumbent's market share will accordingly plummet quickly below whatever market share is necessary to trigger the post-entry price floor, and then the incumbent will be free to

adopt above-cost price cuts that drive the less efficient entrant out of the market.

In markets where there are few physical limitations on entrant expansion, the drop in incumbent market share may be nearly instantaneous. In particular, in the airline industry, which was the genesis of these proposals, airplanes are relatively easy to move when demand increases on some routes, and relatively easy to lease if total demand for the airline rises. There thus may be no effective barrier to an entrant expanding to take all the consumer demand that might respond to its lower prices. In many technology or intangible markets, there may likewise be few physical limitations to expanding entrant market share, as when output expansion merely requires more software downloads.

Even if the less efficient entrant must ramp up its capacity over time, an entrant with a price advantage will sooner or later take enough market share to deprive the incumbent of its monopoly share. It seems likely to be sooner rather than later when one considers four additional points. First, monopolists rarely have a hundred percent market share, but rather normally begin the post-entry period with a market share only somewhat above whatever threshold defines monopoly power. They thus need not lose much market share to lose their monopoly power. Second, as noted above, such post-entry price floors give incumbents incentives to raise prices and *lose* market share as rapidly as possible to bring closer the day when the restriction expires and they can drive the entrant out and restore monopoly pricing. Third, while efficient firms are limited in number, the world of *less* efficient firms is hardly scarce, so that if the restriction encourages entry by any of them, it is likely to encourage entry by lots of them, all of which can ramp up capacity simultaneously.

Fourth, the relevant set of entrants is likely to be in industries that permit rapid expansion. That relevant set consists of those less efficient entrants whose entry might actually be caused by an above-cost floor on post-entry incumbent pricing. That causal link requires two things. (a) These must be entrants that would *not* have entered without the restriction. And that, as noted above, is disproportionately likely to be in industries where capacity can be expanded rapidly, because that permits incumbents to drive out entrants quickly with temporary price cuts. Where capacity cannot be expanded so rapidly, then the restrictions on reactive above-cost price cuts will last longer, but are less likely to have been necessary to encourage the less efficient entrant to enter at all. Thus, the very factor that makes a post-entry price floor likely to encourage less efficient entrants—an industry where capacity can be expanded rapidly—also tends to mean that any post-entry price floor will be very short-lived. (b) It must be the case that the restrictions do provide a meaningful inducement to less efficient entry. Because the restrictions only offer protection for a limited

time (twelve to eighteen months at the outside under the Edlin and Williamson rules), they are unlikely to encourage less efficient firms to enter when entry requires large capital investments that cannot be recouped in a short period. Yet such large capital investments are the major reason why entrants might need time to ramp up capacity. Thus, the set of less efficient entrants whom the restrictions might actually encourage to enter probably did not need large capital investments and are thus more likely to be easily expandable.

In short, while the proposed above-cost floors on incumbent post-entry prices should encourage some additional entry by less efficient firms, that encouragement will be relatively weak because the incremental protection offered by such price floors will be short-lived. Less efficient firms will realize that the restrictions will not enable them to stay in the market in the long run. Thus, they will be encouraged to enter only when this marginal prolongation in the short-run period during which they can profitably remain in the market provides the additional increment necessary to make total short-run profits exceed the sunk costs of entry. Where entry costs are significant, this will be rare, because large entry costs cannot be covered with short-run profits. Where entry costs are small, this is unlikely, given that less efficient entrants will generally not be discouraged by the prospect of reactive above-cost price cuts anyway because they can cover small entry costs with short-term profits.²²⁵ Further, in industries that require so little capital investment, incumbents are unlikely to have any market advantage that makes them more efficient than entrants at all. Thus, the restrictions will encourage entry only when entry costs are in an intermediate range that is large enough to deter less efficient entry given the short-run profits that could be made without the restriction, but not so large to deter entry given the slightly larger short-run profits that could be made with the restriction.

There is an additional reason why the Williamson and short-term profit-maximizing rules would provide weak encouragement to less efficient entry. Namely, these rules set a post-entry price floor that is often too low to prevent the incumbent from driving out the less efficient entrant with an above-cost price cut.²²⁶ In such cases, those rules cannot offer any protection to less efficient entrants that might encourage their entry. Since they will be ineffectual in protecting less efficient entrants, their only post-entry effect in such cases will be the harmful one of sometimes preventing

225. CARLTON & PERLOFF, *supra* note 19, at 337 (noting that predatory strategies cannot succeed against entrants when entry costs are low).

226. See Edlin, *supra* note 6, at 957-59, 977-78, 981-82 (rejecting the short-term profit-maximization test because a price at that level can sometimes drive out less efficient entrants); Williamson, *supra* note 7, at 297-98 (modeling the case where maintaining the incumbent's pre-entry output level does not leave sufficient market output for the entrant to operate at a large enough scale to profit at a lower price).

the incumbent from cutting prices even further (that is, below the price floor), which would have benefited consumer welfare and increased allocative efficiency.²²⁷

Indeed, Professors Ordovery and Willig assumed their short-term profit-maximization test could *never* protect a less efficient entrant.²²⁸ Their reasoning was that if the incumbent priced above entrant costs, it would lose all production to the entrant. Thus, pricing slightly below a less efficient entrant's costs would always be the more profitable alternative. But if the entrant is capacity-constrained over the short run, then the entrant will not be able to take all market output. Instead, the incumbent will be left with a residual demand curve determined by subtracting entrant output from the total market demand curve, and pricing above cost will likely maximize the incumbent's short-run profits. Professors Ordovery and Willig would also apply their test to condemn an above-cost price cut in one product if it diverted sufficient profits from another substitute product made by the incumbent.²²⁹ But if the substitute product enjoys any supracompetitive profit margin, this test would prevent what is effectively an efficient price cut that brings the price on the combination of products closer to their cost, and would protect (quite undesirably) a less efficient entrant in one product to preserve the incumbent's supracompetitive profits in the other product. Still, if the entrant is not capacity-constrained and such substitution effects are irrelevant, Ordovery and Willig appear to be correct that their price floor cannot protect less efficient entrants absent erroneous application. Further, even if an entrant begins with a capacity constraint, eventually its output will rise sufficiently to raise this problem absent substitution effects. This confirms the point above that the short-term profit-maximization test cannot offer any long-term protection to a less efficient entrant. And if the goal is to deny protection to less efficient entrants, a price-cost comparison test will be better because it denies protection to less efficient entrants when capacity constraints or substitution effects matter and is generally easier to apply accurately.²³⁰

Likewise, Williamson also assumes his rule will never encourage entry by a less efficient firm. He reaches this conclusion by assuming that the incumbent always has the knowledge and desire to set pre-entry output sufficiently high that maintaining that output after entry will make entry

227. Any benefits in such cases would instead have to be based on the claim that the rule encourages a pre-entry incumbent-output expansion that amounts to a form of limit pricing. See *infra* Section IV.D.

228. See Ordovery & Willig, *supra* note 71, at 18-19.

229. *Id.* at 20-21.

230. See *supra* Part II; *infra* Part V.

unprofitable.²³¹ But these assumptions about incumbent knowledge and desire are both false. Williamson himself recognizes that in fact no incumbent is that prescient. Instead, there is a range of probabilities so that the incumbent will have to set pre-entry output at an average that will sometimes make entry unprofitable but sometimes will not.²³² Second, his premise that the incumbent will always want to set pre-entry output high enough to make entry unprofitable rests on what Williamson acknowledges is the "arbitrary assumption" that incumbents strictly prefer avoiding post-entry hazards to earning pre-entry profits.²³³ If we instead adopt the more rational assumption that the incumbent attaches some positive value to pre-entry profits, they will make trade-offs that lower pre-entry output somewhat, and will thus sometimes be unable to drive out a less efficient entrant under a rule that prohibits output expansions. Indeed, incumbents would have strong incentives to do so since any increased pre-entry profits will not have the time and uncertainty discount applied to fears of a decline in post-entry profits.²³⁴ Thus, although the encouragement to entry will be weak, the Williamson output ceiling will sometimes prevent the incumbent from being able to drive out a less efficient entrant.

The lack of encouragement to entry by less efficient firms will be even greater if the restriction is rendered ineffective (as discussed below) either by buffer zones established to escape the difficulty of adjusting for demand or cost shifts, or by a failure to regulate nonprice reactions.²³⁵ It will also be even more ineffective if the restriction is defined to begin at a moment of entry that is not sufficiently early (and long-lasting) to restrain reactive price cuts that anticipate entry.²³⁶ Nor can these likely sources of regulatory ineffectiveness be easily avoided since doing so requires incurring the serious adverse effects of mistaken adjustments in price controls, freezing innovation, or a lengthier distortion of prices and innovation.²³⁷

b. *The Effects of (Weakly) Encouraging This Additional Less Efficient Entry*

Since any encouraged less efficient entrant cannot survive in the market once the restriction expires, the restriction cannot have any long-term

231. See Williamson, *supra* note 7, at 294, 297-98, 310 n.66. Williamson assumes that an entrant with the same cost curve as the incumbent will be left at zero profits, which means a less efficient entrant with a higher cost curve would suffer an actual loss.

232. *Id.* at 294 n.33.

233. *Id.* at 314.

234. See *infra* Subsection IV.D.1 (noting other reasons why the incumbent may not keep pre-entry output so high).

235. See *infra* Sections V.B-C.

236. See *infra* Section V.A.

237. See *infra* Part V.

beneficial effect on post-entry market pricing. Further, any short-run effects may be very short indeed given the reasons noted above for thinking the restrictions will expire relatively quickly.

But such restrictions will have one clear positive short-run effect. Namely, those buyers who purchase from the less efficient entrant during this short-run period will pay a lower price than they would have paid if that entrant had not been induced to enter by the restriction.

However, those buyers who continue to purchase from the incumbent during this short-run period may or may not pay a lower price than they otherwise would have. The reasons are several. First, as noted above, incumbent monopolists have incentives to respond strategically to such restrictions by raising their prices to lose market share and speed the day when the restriction expires. The restrictions even give incumbents perverse incentives to raise prices above their short-run profit-maximizing level.²³⁸ Sometimes these perverse incentives can cause the incumbent to increase prices above its pre-entry level. It will not always do so because the addition of entrant output will itself lower the short-term profit-maximizing price of the incumbent by leaving less residual demand for the incumbent (absent an offsetting increase in total market demand). Thus, if the pre-entry price was at the short-term profit-maximizing level, then that same price is (absent a demand increase) likely to be already above the post-entry short-term profit-maximizing level. Whether going even further above that profit-maximizing level will be a cost-effective way to speed the expiration of the restriction will depend on the particular facts.

Second, as also noted above, the restrictions will generally require a reduction from the incumbent's pre-entry output level that can reduce the incumbent's productive efficiency. Increased incumbent costs will thus increase the incumbent's short-term profit-maximizing price. Especially in combination with the fact that the restriction would give the incumbent incentives to raise prices above short-term profit-maximizing levels, this will further increase the likelihood that the incumbent will raise prices above pre-entry levels.

Third, sometimes the incumbent's pre-entry price will reflect an attempted limit price (that the incumbent mistakenly set a bit too high to deter entry) that was below the short-term profit-maximizing level from the start.²³⁹ Because the restriction (in any case where it encouraged entry) makes it impossible to drive out the entrant, the incumbent will have no reason not to raise post-entry prices to at least the short-term profit-maximizing level until the restriction expires. Combined with the likely

238. See *supra* text accompanying note 220.

239. A limit price is an above-cost price deliberately set by a monopolist below its short-term profit-maximizing level in order to preclude entrants. See VISCUSI ET AL., *supra* note 57, at 168-70.

increase in that level because of increased incumbent costs and its incentives to charge over that level to speed the end of the restriction, this means the restrictions are especially likely to increase incumbent prices above pre-entry levels when pre-entry prices reflected attempted limit prices.

Accordingly, buyers who purchase from the encouraged less efficient entrant will pay less post-entry, but those who continue to purchase from the incumbent may pay more. And more buyers will be in the latter camp than the former. The reason is that the restrictions only apply as long as the incumbents have a monopoly or dominant market share. This normally means that the incumbent will have over a 50% market share during the period of any restriction, and that most consumers will accordingly be buying from the incumbent. Indeed, if monopoly power is required, probably the great majority of consumers will be buying from the incumbent, given most definitions of monopoly-share thresholds. To be concrete, suppose that in a particular industry the incumbent has 100% market share and the minimal market share at which a firm will be said to have monopoly power triggering the relevant restriction is 70%. In the short run before the incumbent drops below its monopoly share, the restriction would allow the entrant to ramp up from 0% to 30% market share and cause the incumbent to ramp down from 100% to 70%. If the rate at which the incumbent loses market share is constant, then during this period an average of 85% of buyers will continue to buy from the incumbent.

The fact that most buyers will continue purchasing from the incumbent makes it more likely that the net effects are negative for consumer welfare. In this example, on average only 15% of buyers would benefit from lower entrant prices during this short-run period. If the pre-entry price were \$100, and the entrant priced at \$80 (as it would under Edlin's 20% discount rule), there will be a net harm to consumer welfare if the desire to speed the demise of the entrant causes the incumbent to raise prices to any level above \$103.53.

Further, to avoid making the restrictions entirely ineffectual, they have to begin when entry is first foreseeable, rather than when the entrant first sells, or else the incumbent would just cut prices before entrant sales begin.²⁴⁰ Thus, the period during which incumbent prices will be elevated by the restriction will last longer than the period during which consumers will enjoy lower entrant prices.

In short, even in the case where the restrictions do encourage the entry of a less efficient firm that otherwise would not have occurred, the net effect on post-entry consumer welfare is mixed, and thus so too is the predicted effect on allocative efficiency. Other effects of the restrictions in

240. See *infra* Section V.A.

such a case are unambiguously negative. Productive efficiency will suffer because, by hypothesis, the restriction here has resulted in a shift of market production to a less efficient firm. Further, the reduction in the incumbent's pre-entry output will inflict wasteful transition costs and likely decrease the incumbent's productive efficiency.²⁴¹

Finally, the capital costs of entry will be wasted because the less efficient entrant who was encouraged by the restriction to enter will eventually be driven from the market. Indeed, the less efficient entry encouraged by the restrictions would amount to a form of wasteful rent-seeking. The entrant will be encouraged to expend those capital costs only because of the short-term profits it earns by shifting producer surplus from the more efficient incumbent to itself. As we will see below, this transfer in the rewards for creating a more efficient product or method of production from the firm that created it to another firm that did not will reduce the incentives to invest in such efficient creation.²⁴² But here the point is that some, and at the extreme all, of this transferred producer surplus will be dissipated by the expenditure of entry costs that otherwise would have been avoided.²⁴³ Such dissipation results in an efficiency loss.

B. *Effects on Likelihood and Consequences of Efficient Entry*

Proponents of restrictions on above-cost pricing have analyzed their effects on less efficient entrants and, to a lesser extent, on entrants that are initially less efficient but grow more efficient with time.²⁴⁴ But they have ignored the effects of their proposed restrictions when entrants are just as efficient as, or more efficient than, the incumbent. This one cannot do if one wishes to understand the full effects of these restrictions because equal or greater efficiency characterizes many entrants. After all, long-term prospects of at least remaining in the market, if not besting the incumbent, are normally what motivates entry and persuades capital markets to fund it. Indeed, given the weak encouragement the proposed restrictions would give to less efficient entrants, it would seem that the lion's share of entrants would continue to be efficient even if the restrictions were adopted. And the restrictions do have serious effects on efficient entrants.

To be sure, the restrictions will not affect the likelihood of entry by a created entrant that is at least as efficient as the incumbent. To the contrary, if we have set our cost measure correctly, then by definition such an entrant

241. See *supra* Subsection IV.A.1.b.

242. See *infra* Section IV.D.

243. See Richard A. Posner, *The Social Costs of Monopoly and Regulation*, 83 J. POL. ECON. 807 (1975) (noting that monopoly rents will tend to be dissipated by costly competition over which producer gets those rents, unless that competition has socially valuable by-products).

244. See *supra* Section IV.A; *infra* Section IV.C.

could not have been deterred by the prospect that the incumbent might react to entry with an above-cost price cut.²⁴⁵ We thus need to turn to post-entry effects and the ex ante incentives to create efficient entrants.

1. *Post-Entry Effects*

The restrictions will all limit the post-entry competition that otherwise would have occurred between incumbents and efficient entrants. As in the case of less efficient entrants, here too the price floors, where they have bite, will prevent the incumbent from adopting above-cost price cuts that lower prices as much as they otherwise would have. Likewise, the restrictions will give efficient entrants incentives to price no lower than the incumbent's price floor, and perhaps even to begin with a higher price until they can test where that price floor is, even when pricing down to their own costs would have produced a lower price. Consumers who buy from either the incumbent or the entrant will thus pay higher above-cost prices than they would have paid without the floor on incumbent post-entry above-cost pricing. This will harm both consumer welfare and allocative efficiency. The restrictions can also lower incumbent productive efficiency.

Indeed, for efficient entrants, the effects are even more likely to be harmful for three reasons. First, because more efficient entrants have lower costs, it is more likely they would have otherwise set a price further below this price umbrella. Second, their lower costs mean the prices they otherwise would have set would have been lower. Third, for efficient entrants, these negative offsets are not possibly offset by encouraging additional entry. Instead, the effects on entry are on balance negative.

2. *Ex Ante Effects on Creation of More Efficient Entrants*

While the proposed restrictions would have no positive ex ante effects on entry by efficient entrants, that does not end the ex ante analysis. Rather, we need to take it one further step ex ante, to consider what effects such restrictions have on ex ante incentives to create these more efficient entrants. Entrants who are more efficient than incumbents are not magically generated. They require creative effort and capital investments. Both are scarce. We must thus consider the likely effect the proposed restrictions would have on whether this scarce effort and capital will be allocated to these or other forms of entrants.

245. See *supra* Part II. Even if the reader does not agree with my particular cost measure, the conclusions here follow under whatever definition of costs the reader *does* believe suffices to meet this condition.

To the extent proponents are right that these restrictions do encourage entry, it will be by less efficient entrants.²⁴⁶ Since effort and capital is scarce, this increased entry by less efficient entrants must divert effort or capital that otherwise would have gone to the more efficient entrant. Of course, if investors knew one entrant was less efficient and the other was more efficient, they would all choose the latter. But, in fact, there will often be a probabilistic judgment, where a new firm has, say, fifty percent odds of being more efficient and fifty percent odds of not being more efficient. The proposed restrictions effectively reduce the difference in returns between less efficient and more efficient entrants, and thus at the margins induce more investment in less efficient entrants as compared to others that might be more efficient.

True, any diversion of effort or capital to a less efficient entrant will be offset to the extent that, by increasing the short-run returns to entry, the restrictions increase the total effort and capital that flow to all entrants from other areas of the economy. But if we generally thought that such an indiscriminate subsidy to entrants, whether efficient or not, were warranted, then it would cut far beyond the present context to suggest a general "infant firm" policy of subsidizing entrants.²⁴⁷ Moreover, an additional factor means the restrictions would likely decrease the overall returns to more efficient entrants and further discourage the creation of efficient entrants. Creating an entrant to challenge an incumbent is always risky, and thus a major motive for making such a risky investment will often be the prospect of long-term supracompetitive profits if the entry is successful. In particular, investors often invest to create a more efficient entrant based partly on the prospect that, if the entrant truly is more efficient than the incumbent, it can displace that incumbent and become the new monopolist (with lower costs or better quality) and reap supracompetitive profits itself. But to the extent the restrictions do protect and encourage more entry by less efficient entrants, they will, as proponents stress, reduce the profitability of firms that enjoy monopoly profits as a result of their greater efficiency.²⁴⁸ This effect will reduce the potential upside of making an investment that succeeds in creating a more efficient entrant, and thus will lessen the incentives to make risky investments that are necessary to create more efficient firms at all. This effective reduction in the rewards for improving market efficiency will naturally lead to fewer such improvements, resulting in a loss of productive efficiency and fewer more efficient entrants.

246. See *supra* Section IV.A.

247. See *infra* Subsection IV.C.2.

248. See *infra* Subsection IV.D.2.

Consider the inventor who is deciding whether to devote her time to research that has a 50% chance of resulting in a more desirable or efficient product, and a 50% chance of coming up empty. Or consider the venture capitalist who is deciding whether to make an investment in a new technology that has a 50% chance of being preferred by buyers to the incumbent product, but a 50% chance of flopping. In either case, whether the inventor or venture capitalist makes the necessary investment of time and money will depend on how great the returns are if the product does turn out to be better or cheaper. If the returns of a successful product are higher, they are more likely to make investments that lead to more efficient entrants. If the returns are lower, they are less likely.

In short, to the extent the proposed restrictions succeed in their goal of encouraging less efficient entry and inducing limit pricing on firms that acquire market power, they will tend to lessen the creation of more efficient entrants by diverting some effort and capital and by lessening the long-run return on successfully creating a more efficient entrant. Any reduced creation of efficient entrants will cause unambiguous harm to consumer welfare, allocative efficiency, and productive efficiency since entry by efficient entrants not only undercuts monopoly prices but can actually lower costs or improve product quality. Further, these harms will be permanent and long-term whenever a more efficient entrant is discouraged, as opposed to the possible (mixed) benefits when a less efficient entrant is encouraged, which can only last for the short term before the restriction expires.

Proponents of restrictions on reactive above-cost price cuts have ignored these *ex ante* effects on the creation of more efficient entrants. They tend to assume instead that entrants either have an inherent inefficiency disadvantage, or one that just depends on where their output is located on a cost curve equally available to entrant and incumbent.²⁴⁹ This assumes away competition in making the sorts of innovations and investments that can lower cost curves and raise demand curves. Thus, though these models pride themselves on taking dynamic account of strategic intertemporal considerations,²⁵⁰ and do improve on prior static models in that regard, they end up being very static in their assumptions about where the cost and demand curves lie, and they ignore the dynamic possibility that those curves might be changed by innovation or investment. If one instead takes those dynamic effects into account, the effects of the restrictions become even more negative.

249. See, e.g., Edlin, *supra* note 6, at 955-60, 973-78; Williamson, *supra* note 7, at 295, 297-98, 313.

250. See, e.g., Williamson, *supra* note 7, at 284.

3. *The Restrictions Cannot Reasonably Be Construed or Modified To Eliminate Their Adverse Effects on Efficient Entrants*

Although the other restrictions on their face apply regardless of the efficiency of the entrant, one might think Edlin has avoided adverse effects on efficient entrants by providing that his ban only applies in cases where the “incumbent monopoly enjoys significant advantages over potential entrants.”²⁵¹ This sounds like it excludes any protection for more efficient entrants, and perhaps even for entrants whose efficiency disadvantage is small enough to make it plausible that they will overcome it. But Edlin later disavows any such limitation, arguing for a ban on reactive price cuts that applies to any incumbent monopoly.²⁵²

Still, one might be tempted to modify any of the restrictions on reactive above-cost price cuts by defining them to exclude cases where entrants are not initially less efficient. But such modifications would face numerous difficulties.

First, neither regulators nor antitrust litigation would seem likely to gauge accurately when an entrant is less efficient than an incumbent. A cost-based predation test allows market pricing and competition to sort out the efficient entrants from the inefficient ones naturally. But if regulators or antitrust litigation were to apply a cost-based test to efficient entrants and an above-cost price floor for inefficient entrants, then they would have to make freestanding assessments of the efficiency of an entrant. This would be a difficult assessment to make, especially since entrants would have incentives to pretend to be less efficient than they really are in order to get the benefit of a price umbrella. To the extent regulators or litigation erroneously determined entrants were less efficient when they were not, or firms predicted they would err, the effects would be unambiguously adverse.

Second, any modification that resulted in a rule whose substance differed depending on whether a regulator or antitrust litigation would conclude that the entrant was less efficient would violate fundamental rule-of-law norms of providing notice to incumbents about how to conform their behavior to legal dictates. A cost-based test may be complex but provides some notice. The incumbent has access to information about its own costs and prices and by considering it can conform its behavior to the law. But a rule whereby the cost-based test did not apply when the entrant was less efficient would make the substantive rule turn on information about the entrant that the incumbent may not know. Indeed, it might well be that it is

251. Edlin, *supra* note 6, at 945.

252. *Id.* at 967-68.

only by making above-cost price cuts that an incumbent could reveal whether or not the entrant were less efficient. But once it has made such price cuts, it might discover from the market effect on the entrant that its price cut was illegal. Not only would this violate rule-of-law norms, it would lead risk-averse incumbents to avoid above-cost price cuts whenever the relative efficiency of the entrant seemed at all ambiguous.

Third, even if an exception for efficient entrants could be applied with perfect accuracy and predictability, this would not eliminate—and indeed would exacerbate—the adverse effect of the restrictions on the creation of efficient entrants. It would not eliminate this adverse effect because it would flow from the application of the restrictions to less efficient entrants, which at the margin would divert some effort and capital away from the creation of more efficient entrants and reduce the long-run returns to creating a more efficient entrant. It would instead exacerbate this adverse effect because it would give less efficient entrants an extra return from a short-term price umbrella that would be unavailable to more efficient entrants, thus increasing the tendency of the restriction to divert effort and capital from more efficient entrants to less efficient entrants. Indeed, entrants might even have incentives to lower, at least temporarily, their efficiency to try to gain the advantage of such a price umbrella. This would only worsen the effects of the restriction.

C. *Effects for Entrants Who Can Overcome Their Initial Efficiency Disadvantage*

Although the proposals to restrict reactive above-cost price cuts have mainly been based on the premise that the entrant has an insurmountable efficiency disadvantage, they have also sometimes cited the hope that over time the entrant can overcome this efficiency disadvantage.²⁵³ There are two reasons this might happen: The efficiency of the entrant might increase, or the efficiency of the incumbent might deteriorate. Those advocating the restrictions have emphasized the former. And it has some basis. There might be economies of scale or scope at low output that are only available over time as production is ramped up. Or experience in the industry might lower costs or improve quality. But, as noted above, there are also various reasons incumbent efficiency might deteriorate when faced with a less efficient entrant under the proposed restrictions.²⁵⁴ We must thus consider both of these reasons why an entrant might with time be able to overcome an incumbent's initial efficiency advantage.

253. See *id.* at 975 & n.95, 977; Williamson, *supra* note 7, at 296, 298 n.43, 303-04, 313.

254. See *supra* Subsection IV.A.1.b.

1. *When Overcoming Incumbent Efficiency Advantage Necessitates Some Deterioration in Incumbent Efficiency*

One possibility is that any increased entrant efficiency does not alone suffice to overcome the initial efficiency disadvantage, but that the combination of any increased entrant efficiency with the deterioration in incumbent efficiency caused by the restriction *does* suffice. Because in these cases the degradation of incumbent efficiency was necessary to overcome the initial efficiency advantage, the final efficiency of both the entrant and incumbent must be less than the initial efficiency of the incumbent.²⁵⁵

Such entry will indeed be encouraged by the restriction, for without the restriction the entrant never would have been able to compete effectively with the incumbent. The incumbent would just have lowered its post-entry price to an above-cost level that enabled it to maintain output and fend off any efficiency degradation, while still undercutting the entrant and driving it out of the market.

But the consequences of encouraging such entry are likely to be undesirable. It effectively changes an efficient monopoly market into a market with two or more inefficient firms. It is highly unlikely that this will be socially desirable. First, the two inefficient firms may engage in duopoly pricing that is just as supracompetitive as monopoly pricing but that, because costs are higher, means higher prices.

Second, even if enough other less efficient firms entered to make the market competitive, productive efficiency generally matters much more than allocative efficiency. Even in static models, the efficiency gains from a small cost reduction usually offset the efficiency loss from a large price increase.²⁵⁶ The basic reason is that the cost reduction creates efficiency gains for all output, whereas the price increase produces an efficiency loss only for the marginal reduction in output. True, it is a disputed issue whether antitrust law does (or should) protect just consumer surplus or total

255. In the extreme, some of these cases will be ones where the entrant experienced no efficiency improvement but was able to overcome its initial efficiency disadvantage solely because of deteriorating incumbent efficiency.

256. Professor Williamson has shown that even at a very high demand elasticity of 2, a cost decrease of 0.25% offsets a price increase of 5%, and a cost decrease of 9% offsets a price increase of 30%. See Oliver E. Williamson, *Economies as an Antitrust Defense: The Welfare Tradeoffs*, 58 AM. ECON. REV. 18, 22-23 (1968). At a more normal demand elasticity of 1, it takes half the cost decrease to offset the same price increases: A 0.12% cost decrease offsets a 5% price increase, and a 4.5% cost decrease offsets a 30% price increase. *Id.* At a lower demand elasticity of 0.5, a 0.06% cost decrease offsets a 5% price increase, and a 2.25% cost decrease offsets a 30% price increase. *Id.*

efficiency (measured by the sum of consumer and producer surplus).²⁵⁷ Judge Robert Bork's argument for the latter was roundly critiqued as based on the premise that monopolists are owned by shareholders who are consumers too.²⁵⁸ But Bork's proposition seems more distributionally attractive now that most workers are invested in stocks through their pension plans. More persuasively, one might add that the per capita income of any nation must in the end rest on its productivity. More productive efficiency thus generally means higher wages for workers. Accordingly, increases in productive efficiency benefit consumers both as employees and investors, making it more likely that consumers will be better off when the productive efficiency gain outweighs the loss in consumer surplus. The odds increase even further when one takes into account that any increased productive efficiency will also increase tax receipts that benefit the general citizenry. Indeed, some argue that taxes can generally achieve any redistributive aim better than substantive law, and that therefore substantive law should focus solely on wealth-maximizing efficiency and leave the redistribution to taxation.²⁵⁹

Third, increased productive efficiency may eliminate any harm to consumer welfare because lower costs tend to offset any tendency of monopolies to increase prices.²⁶⁰ That might be one reason the evidence turns out to be quite disputed about the degree to which high market shares even produce higher prices.²⁶¹ Some conclude that the degree to which market shares fluctuate influences market performance far more than the size of market shares.²⁶² Professor Richard Schmalensee's review of the

257. Compare Robert H. Lande, *Wealth Transfers as the Original and Primary Concern of Antitrust*, 34 HASTINGS L.J. 65, 68-70 (1982), with ROBERT H. BORK, *THE ANTITRUST PARADOX* 107-15 (1978).

258. BORK, *supra* note 257, at 110.

259. See Louis Kaplow & Steven Shavell, *Why the Legal System Is Less Efficient than the Income Tax in Redistributing Income*, 23 J. LEGAL STUD. 667 (1994).

260. The monopolist price will reflect a markup over cost that depends on demand elasticity. See, e.g., PINDYCK & RUBINFELD, *supra* note 40, at 339. Thus, any reduction in cost reduces the difference between the monopoly and competitive prices. If the monopoly has sufficiently lower costs, the monopoly price will be less than the competitive price. See, e.g., 1992 Horizontal Merger Guidelines, 57 Fed. Reg. 41,552 § 4 (Sept. 10, 1992) (establishing that a merger that creates market power has a defense if it creates increased efficiencies that fully offset the tendency of the market power to increase prices).

261. See CARLTON & PERLOFF, *supra* note 19, at 258-59; Harold Demsetz, *Industry Structure, Market Rivalry, and Public Policy*, 16 J.L. & ECON. 1 (1973); Harold Demsetz, *Two Systems of Belief About Monopoly*, in INDUSTRIAL CONCENTRATION: THE NEW LEARNING 164 (Harvey J. Goldschmid et al. eds., 1974); Sam Peltzman, *The Gains and Losses from Industrial Concentration*, 20 J.L. & ECON. 229 (1977); Ky P. Ewing, Jr., *The Soft Underbelly of Antitrust: Some Challenging Thoughts for the New Millennium*, ANTITRUST REP., Sept. 1999, at 2; Barry C. Harris & David D. Smith, *The Merger Guidelines v. Economics: A Survey of Economic Studies*, ANTITRUST REP., Sept. 1999, at 23.

262. See R.E. Caves & M.E. Porter, *Market Structure, Oligopoly, and Stability of Market Shares*, 26 J. INDUS. ECON. 289 (1978); Mariko Sakakibara & Michael E. Porter, *Competing at Home To Win Abroad: Evidence from Japanese Industry*, 83 REV. ECON. & STAT. 310, 312 (2001).

literature concludes that while the “relation, if any, between seller concentration and profitability is weak statistically” in studies comparing the concentrations in different industries, “[i]n cross-section comparisons involving markets in the same industry, seller concentration is positively related to the level of price.”²⁶³ Since efficiencies are more likely to differ between industries than within the same industry for a firm operating in different geographic markets, this observation is consistent with the conclusion that concentration earned by greater efficiency generally does not increase prices, but concentration produced by other (nonmerit) factors does. Since here the initial incumbent is (by hypothesis) more efficient, there is little reason to think its replacement with less efficient firms would benefit buyers with lower prices even in the short run.

Fourth, if one moves away from static models to dynamic ones, it is clear that in the long run the pace of innovation advances consumer welfare far more than maintaining allocative efficiency.²⁶⁴ Indeed, it has been shown that nations with better market performance generally compete by innovation and differentiation rather than by price and imitation.²⁶⁵ Replacing an efficient incumbent with less efficient firms would reverse this dynamic process. And those less efficient firms are likely to be less innovative. Schumpeter goes even further to argue that firms with higher market power are more likely to innovate because they can reap a larger share of the benefits of their innovation, whereas if there were perfect competition no one would have the incentives to invest in unpatentable product improvements.²⁶⁶ Whether or not that is generally true, it certainly seems likely when the firm with market power is (by hypothesis) more efficient.

263. Richard Schmalensee, *Inter-Industry Studies of Structure and Performance*, in 2 HANDBOOK OF INDUSTRIAL ORGANIZATION, *supra* note 133, at 951, 988.

264. 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 720a, at 255 & n.3 (collecting sources); AREEDA & KAPLOW, *supra* note 132, at 31; JOSEPH A. SCHUMPETER, CAPITALISM, SOCIALISM, AND DEMOCRACY 84-92, 99-106 (3d ed. 1950); *see also* Moses Abramovitz, *Resource and Output Trends in the United States Since 1870*, 46 AM. ECON. REV. 5 (1956); Robert M. Solow, *A Contribution to the Theory of Economic Growth*, 70 Q.J. ECON. 65 (1956); Robert M. Solow, *Technical Change and the Aggregate Production Function*, 39 REV. ECON. & STAT. 312 (1957).

265. Michael E. Porter, *The Current Competitiveness Index: Measuring the Microeconomic Foundations of Prosperity*, in THE GLOBAL COMPETITIVENESS REPORT 40, 45 (2000).

266. SCHUMPETER, *supra* note 264, at 87-92, 99-106. Professors Areeda and Kaplow have disputed this hypothesis with evidence that firms with market power do not obtain more patents or spend more on research and development. AREEDA & KAPLOW, *supra* note 132, at 31-33. But this misapprehends Schumpeter's point, which was that innovation includes not just technological progress but changes in organization, distribution, or scale that are *not* protected by patents and would thus go unrewarded without some degree of market power. SCHUMPETER, *supra* note 264, at 84-85, 88-89. The huge investments necessary to create hub-and-spoke airline systems would be just such an example. *See supra* Section III.A. Indeed, properly understood, Schumpeter's theory would predict firms that lack market power would have greater incentives to shift their innovation investments toward research and development designed to obtain patents because that is the only form of innovation for which they can exclude competition and obtain rewards.

In sum, when entrants can overcome their initial efficiency disadvantage only if the restrictions somewhat lessen the incumbent's efficiency, then the effects on prices and thus on consumer welfare are ambiguous, but the effects on productive efficiency are clearly negative.

2. *When Increased Entrant Efficiency Suffices To Overcome Incumbent Efficiency Advantage*

The other possibility is that entrant efficiencies alone will increase sufficiently with time to overcome its initial efficiency disadvantage before the restriction on reactive price cuts expires. This can include cases where incumbent efficiency declines, as long as the final efficiency of the entrant exceeds or equals the initial efficiency of the incumbent. This might be true when there are economies of scale the entrant can only access over time, and the minimum efficient scale is less than half the potential market output. It might also happen when the increased entrant efficiencies come from experience and learning by doing, which need not come at the expense of incumbent efficiency.

But the analysis that follows shows such entrants do not need encouragement from a restriction on reactive above-cost price cuts. Thus, the effects of a restriction in their case are undesirable.

a. *Why Such Entrants Would Generally Enter Without Any Restriction on Reactive Above-Cost Price Cuts*

If it is possible to ascertain when increased entrant efficiency alone will allow it to overcome the incumbent's initial efficiency advantages, then such entrants should be able to persuade capital markets to lend them enough money to get established without the protection of any restriction on reactive above-cost price cuts. True, the initially less efficient entrant will suffer start-up losses if the incumbent's above-cost price is below the entrant's initial costs. This can force the entrant to charge a below-cost price to remain in the market, which is one more reason to allow entrants to charge promotional prices.²⁶⁷ But this initial need for a fund to cover start-up losses will simply be one of the many capital costs of entry that must be considered and that any entrant would anticipate.

When the initial inefficiency results from inexperience, the investment will effectively be in human capital, the cost of which requires funding the losses necessary to get that experience. But there is no reason to treat investments in the human capital necessary to enter a market successfully any differently from investments in the physical capital necessary for

267. See *infra* Subsection IV.C.3.

successful entry. Nor is this entry cost an artificial one, since it reflects a real societal loss of efficiency from shifting production in initial stages to a less efficient firm. If the entrant cannot earn a sufficient return to cover this entry cost, there is no more reason to think its entry desirable than it would be for an entrant that cannot cover the capital cost of building a plant. Thus, if entry by an initially less efficient firm is itself efficient and desirable, the capital markets should be willing to provide the necessary capital to cover start-up losses, just like they cover other capitalized entry costs.²⁶⁸

One might think that capital markets would not cover these start-up costs because the long-term result will be competition between the entrant and incumbent with both pricing at long-run costs. But here cost has the economic definition that includes a normal rate of return on capital investment (including human capital), and if entry is efficient, that normal rate of return should suffice at a minimum to recoup this and other entry costs. Indeed, as soon as it realizes the entrant cannot be driven out, the incumbent monopolist will have incentives to accommodate entry by pricing at supracompetitive duopoly levels,²⁶⁹ which should more than suffice to cover the entrant's capital investment of bearing the initial inefficiency loss. Accordingly, once the initially less efficient entrant raises the capital to cover initial inefficiency losses, that itself should assure that the incumbent's reactive pricing will eventually be high enough that those initial losses will at least be recouped and may even be immediately high enough that those initial losses will never be incurred.²⁷⁰ If the entrant anticipates eventually becoming more efficient than the incumbent, then it will even be able to drive out the incumbent and reap its own monopoly profits, thus amply covering these start-up entry costs.

Another concern might be the general concern about *any* sunk entry cost—that the threat of the incumbent pricing at its variable costs will suffice to deter entry by an equally efficient entrant. But since this is a general problem, there is no reason to adopt a special doctrine to deal with those sunk entry costs that happen to take the form of initial inefficiency losses due to inexperience. Instead, a more general doctrine must be developed to deal with this issue. For reasons analyzed above, it turns out that an incumbent threat to price at whatever costs are variable to it during any pricing period will not suffice to deter an equally efficient entrant

268. Below-cost incumbent pricing is a different story because it might mislead prospective entrants or capital markets into thinking incumbent efficiencies are greater (or market conditions are worse) than they actually are. See Bolton et al., *supra* note 15, at 2247-49, 2285-330 (synthesizing the recent literature).

269. See *supra* Section I.E.

270. They will be immediately that high when the incumbent shares the belief of the capital markets that the entrant will eventually be as efficient. The incumbent will have incentives to be as accurate as it can in such predictions since, if the entrant will eventually be as efficient, immediately higher prices will maximize the incumbent's profits.

because, once the sunk entry costs are incurred, such pricing cannot drive out the entrant and would thus be irrational.²⁷¹ Even if one did not accept this reasoning, the solution would not be to replace a cost-based test with a restriction on above-cost price cuts. Instead, the solution would simply be to define the cost measure to include the magnitude of the predator's costs for the sorts of sunk entry costs that are variable to the entrant when it decides whether to enter.²⁷² The start-up sunk costs in human or physical capital necessary to achieve equal efficiency with the incumbent would be included in the sort of costs variable to the entrant. But the magnitude of those costs to the incumbent must be determined in a future-oriented way.²⁷³ And since these are start-up losses, they are nonrecurring by definition and thus the incumbent will not face these costs in the future. That is, incurring start-up losses to replace the incumbent's output with the entrant's involves a real efficiency loss if the remainder of their future costs is really identical. Unless the entrant can cover those costs from market returns, then its entry will not really be efficient.²⁷⁴

Thus, entrants who in the long run will be just as efficient as the incumbent should enter without any restriction on reactive above-cost price cuts. The only reason to expect any difference would be if courts are somehow better than firms and capital markets at identifying entrants who have this characteristic, which is surely implausible. Not only do capital markets have far more expertise on this matter, they have a lot more incentive to make correct decisions. Indeed, whether or not they on average are better at identifying good entrants, the capital markets will drive those who prove to be bad at making this identification out of the market, leaving only those who do better.

Williamson points out that capital markets might wrongly fail to provide funds because it is too costly for entrants to disclose their actual state of competitiveness persuasively to potential investors.²⁷⁵ But that information cost is a real societal cost of entry, and, absent more precise information, the capital markets should rely on the average competitiveness of such an entrant, which they can gauge at least as accurately as courts and juries. Williamson's rule (and the other restrictions) would effectively protect *all* entrants without incurring the cost of becoming any more informed. This will induce the entry of some firms that prove to be

271. See *supra* Section II.E.

272. See *supra* Section II.D.

273. See *supra* Section II.D.

274. If one also rejected this future-oriented conclusion, the solution still would not be to restrict above-cost price cuts. It would instead be to define the cost measure to include the amortized cost of the sorts of sunk entry costs that are variable to the entrant when it decides whether to enter. See *supra* Section II.D. That measure would then include any start-up sunk costs in human or physical capital necessary to achieve equal efficiency with the incumbent.

275. Williamson, *supra* note 7, at 304 n.58.

competitive, but will also induce the entry of many firms that are not, and on average will induce more of the latter entry than the former in any case where capital markets were not willing to make the investment given the average competitiveness of the class of entrants.

To put it another way, one could accurately characterize the various restrictions on reactive above-cost price cuts as mandatory consumer financing of the new entrant. Instead of having the financing provided by capital markets, the financing is provided by consumers in the form of higher post-entry prices. And instead of having the financing decision made voluntarily by experts on capital markets, it would be made involuntarily, based either on a regulatory or litigation assessment of the particular entrant or on a mistaken blanket rule that includes all entrants. Indeed, the Edlin and Williamson proposals include a twelve- to eighteen-month limit only as a rule specification of the more general standard that the period of price restriction should last long enough to allow the entrant "sufficient time to recover its entry costs and become viable."²⁷⁶ But the persons from whom the entrant is "recovering" its entry costs will be the consumers who are paying higher post-entry prices than they otherwise would have. And unless there is a (mistaken) global judgment that all entrants can do so, the person making the judgment whether the entrant will become viable (i.e., efficient in the long run) will be the regulator, judge, or jury. There is no reason to think it desirable to have such government-ordered consumer financing of entrants that cannot get financing on capital markets.

If there were good reasons to think capital markets were so imperfect that mandatory consumer financing were desirable, there would be no reason to limit that proposition to the particular set of cases where entrants face incumbents with market power who are likely to drive them out with reactive above-cost price cuts. The proposition would justify protecting all entrants with government subsidization, tariffs, or post-entry price floors and output ceilings. If such "infant firm" arguments for protecting entrants are not persuasive generally, there is no greater reason to find them persuasive here.

In short, any entrant who is likely to experience a sufficient efficiency improvement to overcome an initial efficiency disadvantage will likely get the financing to enter without any restriction and thus cannot have its entry encouraged by the restriction. The restriction is likely to encourage entry only in cases where the government and capital markets diverge in their prediction of whether an entrant's efficiency will rise enough to overcome its initial inefficiency. And the most likely reason for such divergence is that the government has erroneously overestimated the ability of the particular entrant's efficiency to rise or has erroneously overincluded all

276. Edlin, *supra* note 6, at 969; *see also* Williamson, *supra* note 7, at 296.

entrants. Thus, the restriction is more likely to encourage entry by firms that in fact will never overcome the efficiency disadvantage than to encourage entry by firms that will.

Further, even in such cases, the encouragement to entry will be weak. This is because the capital markets by hypothesis will regard the entrant as permanently less efficient and thus unlikely to survive in the market in the long run. Accordingly, the restriction would only cause the capital markets to fund the entrant in the rare case where the anticipated additional profits during the relatively short period of any restriction on reactive above-cost price cuts provide the marginal difference necessary to overcome capital entry costs.²⁷⁷

b. *The Undesirable Consequences*

In short, if capital markets are reasonably efficient, entrants who will become with time as efficient as the incumbent should enter even without any restriction. The effect of a restriction on such entrants will thus be adverse. During whatever initial period the entrant remains less efficient than the incumbent, the effects will be much the same as those described above for less efficient entrants who would enter without any restriction.²⁷⁸ Consumers who purchase from either the incumbent or entrant will pay higher prices. Where the restrictions have bite, increased production will be shifted to a less efficient producer, and the incumbent will suffer a decreased output that lowers its own efficiency. Thus, both consumer welfare and productive efficiency will suffer. After the entrant becomes just as efficient as (or more efficient than) the incumbent, the effects will be those described for the application of the restrictions to efficient entrants.²⁷⁹ Consumers will pay higher prices, reducing consumer welfare and producing a lower output harmful to allocative efficiency.

3. *Entrants That Share the Incumbent's Declining Cost Curve*

One important case to consider is where entrants have the same cost curve as the incumbent, but the curve is declining, so that costs are higher at low firm outputs than at high firm outputs. Not only is the case a recurring one, but it forms the centerpiece of Williamson's famous model, which merits special attention because it purports to prove that a rule prohibiting output expansions in reaction to entry will *always* have favorable welfare

277. See generally *supra* Subsection IV.A.2.a (explaining why restrictions only provide weak encouragement to less efficient entrants).

278. See *supra* Subsection IV.A.1.b.

279. See *supra* Subsection IV.B.1.

effects. In particular, Williamson concludes that his output limitation rule has no effect on post-entry price or output when the entrant's efficiency would (with increased output) increase to match the incumbent's.²⁸⁰ Rather, he concludes that the only post-entry effect of his rule will be to lower incumbent costs because his output ceiling bars the incumbent from reacting to entry by exceeding its optimal plant output.²⁸¹

But his model depends on various questionable assumptions. Williamson reaches the conclusion that post-entry prices and outputs will be unchanged because he assumes that—no matter what the rule—the incumbent will do the same thing post-entry: produce the level of output that, given an assumed categorical consumer preference for the incumbent, leaves an entrant selling at the same price with a low residual output where it has high costs and cannot earn profits.²⁸² His assumptions about incumbent behavior depend heavily on his premises that incumbents have perfect knowledge about the cost curves of potential entrants, that all entrants have identical cost curves, and that the incumbent does not care about pre-entry profits at all and thus picks whatever pre-entry output level minimizes post-entry hazards.²⁸³ Since those assumptions in fact are not true, the incumbent's actual pre-entry output will reflect average expectations and profit trade-offs, and thus an output-limitation rule will sometimes set an effective price floor that prevents decreases in price and increases in output.

More importantly, Williamson's conclusions depend on the critical assumption that, if the entrant and incumbent have identical prices, the incumbent will be able to sell all its output first, leaving the entrant with only the residual demand.²⁸⁴ This is the necessary premise for his conclusion that, even if the entire cost-output curve is immediately available to the entrant, the incumbent will nonetheless (under any rule) be able to set an output that leaves the entrant at the high-cost portion of the curve.²⁸⁵ That is, if we refer to his graph, reproduced as Figure 2, this assumption is what allows him to conclude that the incumbent will choose a

280. See Williamson, *supra* note 7, at 309-10.

281. Any increase or decrease in incumbent output necessarily increases short-run incumbent costs on Williamson's reasonable assumption that incumbent plant size was set to minimize the short-run costs of making the pre-entry output. See *id.* at 297, 300-02, 309-10. Thus, if current law allows an incumbent to increase short-run output in response to entry, it necessarily increases incumbent short-run costs compared to the pre-entry period. But current law only increases firm costs compared to the Williamson rule on his further assumptions that incumbent output will be unchanged under his rule and that entrant output is the same under any rule. Those assumptions are dubious for reasons noted previously. See *supra* text accompanying notes 221-222, 230-234.

282. See Williamson, *supra* note 7, at 294, 295 n.35, 297-98, 310 n.66, 314.

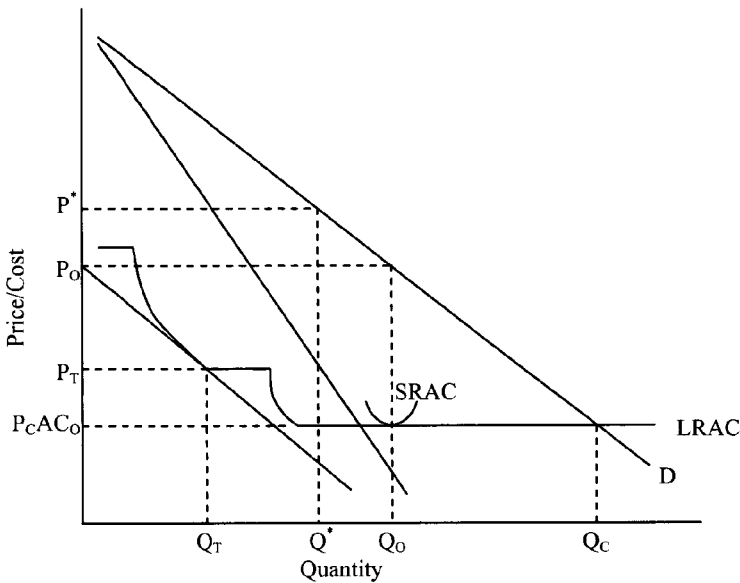
283. See *supra* text accompanying notes 221-222, 230-234.

284. Williamson, *supra* note 7, at 294, 295 n.35, 297-98, 310 n.66.

285. *Id.* at 295, 297-98, 313.

price P_T at which the incumbent will sell all its initial output Q_0 , leaving the entrant with only the residual market output Q_T and thus higher costs.²⁸⁶

FIGURE 2.



But if, as Williamson assumes, the full cost curves were really equally available to the entrant, one could have equally adopted the opposite assumption that the entrant sells all the output it can at an equal price, leaving the incumbent with the residual demand and thus at the low-output, high-cost portion of the curve. That is, Williamson provides no reason to think that at equal price P_T the entrant would not instead produce output Q_0 and leave the incumbent at Q_T . After all, dominant-firm models typically make such an assumption when they assume the dominant firm faces a residual demand curve determined by subtracting the output of the fringe firms at any given price.²⁸⁷

Indeed, in this context, there are good reasons to make such an assumption because buyers would all have an affirmative interest in making sure that the entrant stays in the market, and at an equal price buyers would suffer no individual detriment from dealing with the entrant that might

286. *Id.* at 297, 310 n.66.

287. See CARLTON & PERLOFF, *supra* note 19, at 113-15.

create collective action problems for them.²⁸⁸ If long-term contracting is possible, the entrant with the same cost curve should indeed be able to lock up a sufficient share of the market to put it at least at the minimum efficient scale.²⁸⁹ But even if long-term contracting is not possible, each buyer on the spot market has an incentive to deal with the entrant at any equal price to keep the competition going. If so, the entrant would also reach the low-cost portion of its cost curve and would instead fit the profile of an equally efficient entrant.

Perhaps Williamson is implicitly assuming that the incumbent has a brand-name advantage or familiarity that will lead buyers to choose it at the same price. But if that is the case, it means that, according to the buyers' revealed preferences, the incumbent good is actually more valuable than the entrant's. Because brand-name advantages may not have any concrete manifestation in product quality, they are sometimes dismissed as insubstantial. However, if people are willing to pay more for certain brand names, that means they value the greater predictability and peace of mind that comes with that choice. That is one reason they prefer to buy at McDonald's rather than the unknown hamburger joint. We have no warrant for second-guessing what consumers choose to value, and thus no more reason to question their preference for brand names than to question their preference for vanilla ice cream over pistachio. The revealed preferences of buyers show that brand-name goods are of higher quality in the only sense that is meaningful on a market: Consumers are willing to pay more for it. Thus, if such a brand-name advantage exists, then the two firms either have different demand curves or one must adjust their cost curves to take into account the fact that the cost of producing an equally valued product is higher for the entrant. Either way, Williamson's model would no longer hold. Instead, we would have an entrant who is (at least initially) less efficient at every output level.

Or, rather than adopting either extreme assumption, one could assume buyers have no categorical preference for either the incumbent or the entrant, but would buy from them equally if the price were equal. If that were the case, the entrant could respond to any above-cost incumbent price with a lower entrant price, expanding total market output until half of that output put the entrant on the flat portion of the cost curve. That is, if the incumbent tried to drive out the entrant by selling at P_T as Williamson posits, the entrant would be able to sell at the same price half of total market output, or around Q^* , which, given Williamson's particular drawing, would put the entrant beyond its minimum efficient scale and make entrant

288. Cf. *supra* Subsection IV.A.1.a (noting that collective action problems would be raised if the incumbent could offer a lower price than a less efficient entrant).

289. See *supra* Subsection IV.A.1.a.

sales profitable.²⁹⁰ If the incumbent tried to respond by undercutting that price, the entrant could keep matching or undercutting the incumbent price all the way down to P_C , below which the incumbent could not go without pricing below cost. At that price, the entrant would capture an output of $\frac{1}{2}Q_C$, which by definition will be on the low, flat portion of the cost curve for any drawing that describes a market where the minimum efficient scale is below half of total market output at that price.

In a sense, this is a product of how Williamson drew his demand and cost curves because in his drawing a price equal to the minimum long-run cost produces a total market output that is more than double the minimum efficient scale where an individual firm can enjoy that cost minimum. But this is no graphical artifact because if his curves were *not* drawn that way, then the market would be a natural monopoly because only one firm could stay in the market at the minimum efficient scale. And if the market is a natural monopoly, there can be no successful competition between the entrant and incumbent in the long run. Instead, the situation would be that described above of an initially less efficient entrant that can never become as efficient as the incumbent even though its entry degrades incumbent efficiency and raises entrant efficiency.²⁹¹ For any market that *does* have the sort of demand and cost curves that Williamson posits, an entrant with immediate access to the same declining cost curve as the incumbent is not really an initially less efficient firm at all, but rather a firm that is equally efficient from the beginning, and thus could not be deterred by any incumbent price at or above cost.

Finally, even if one thought customers did have a generic preference for incumbents at the same price, an entrant with the same cost curve as the incumbent could overcome that because, unlike the incumbent, the entrant can offer a promotional below-cost price.²⁹² The entrant need only make the small additional investment of offering a promotional price slightly below cost, which the incumbent could not match since it is constrained to price at cost. The small price advantage will bring enough sales to the entrant to bring its production to the minimum efficient scale.²⁹³ This is a powerful

290. See Williamson, *supra* note 7, at 297.

291. See *supra* Section IV.A.

292. See 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 746a, at 492-95 (noting that the promotional price defense is available only when a firm lacks market power).

293. Using Williamson's model, the entrant would offer a price of $P_C - \epsilon$ (where ϵ is whatever small discount is necessary to overcome consumer inertia to choose an entrant product over an equally valuable incumbent product). Williamson, *supra* note 7, at 297. Given how Williamson draws his model, a one-penny discount would suffice. At a promotional price, the entrant could sell all the output it wanted. But presumably the entrant would stop once it got to an output above the minimum efficient scale since it loses some (albeit small) amount on any sales past that point, and would no longer be able to offer a promotional price if its output got so large that it would be deemed to have enough market power to make the ban on below-cost predatory pricing apply to it. See 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 746a, at 494-95. Thus, once the promotional price has brought the entrant to the minimum efficient scale, the entrant will raise

justification for allowing entrants to offer promotional prices, but provides no justification for condemning above-cost incumbent prices. Indeed, it confirms the conclusion that cases fitting the Williamson model are (even with his assumption of a categorical consumer preference for incumbent output at the same price) effectively the same as the case of an equally efficient entrant described above.

Thus, if one assumes *either* that incumbents do not enjoy a categorical consumer preference at an equal price *or* that any preference can be overcome with promotional pricing, then the growth of an entrant that has the same declining cost curve as the incumbent cannot be contained by just lowering incumbent prices to a level that leaves the incumbent above-cost and the entrant on the high-cost part of the curve. And if incumbents cannot set a price that either drives out the entrant or contains its growth, then the incumbent's incentives will instead lead it to raise post-entry prices in order to maximize short-term profits. This is another reason the Williamson rule in fact would produce an average decline in post-entry output, with corresponding ill effects that include a likely increase in incumbent costs.

In short, if economies of scale and scope are equally available to both entrant and incumbent from the moment of entry, so that they both have the same declining cost curve, there are two possibilities. If the minimum efficient scale is below half the maximum market output, the case is actually one where the incumbent and entrant are equally efficient from the beginning, and the Williamson rule will have all the adverse effects described for such entrants.²⁹⁴ If the minimum efficient scale is higher than half the maximum market output, we have a natural monopoly, with no possibility of long-term competition. The entrant efficiency will increase with its increasing output, but never to the level of the incumbent. The case will thus have all the adverse effects described for a less efficient entrant, with the additional adverse effects that flow from the deteriorating incumbent efficiency.²⁹⁵

prices to cost and both the entrant and the incumbent will be competing with the same costs in the same market.

294. *See supra* Section IV.B.

295. *See supra* Section IV.A. A similar analysis applies if both the entrant and incumbent have economies of scale available from the moment of entry, but their cost curves differ so that at high output one has lower costs than the other. If it is the entrant that has lower costs at high output, it has no need of protection from a ban on above-cost predatory pricing. Without any such ban, the entrant could have entered at a price below the lowest cost of the incumbent and taken over all market output. All the ban can do is raise incumbent prices in the meantime, and perhaps entrant prices, too. The long-run effect will be unchanged—an entrant monopoly—because this is the case of a more efficient entrant. If it is the incumbent that has lower costs at high output, protecting the entrant cannot help in the long run. Even though the entrant's efficiency increases with its growing output, it will not increase to a level that matches incumbent efficiency. Whenever the restriction expires, the incumbent will just lower its price to match its lower cost at high output, drive out the less efficient entrant, and take over the market. In the short run, there will be all the adverse effects described above for less efficient entrants who decrease incumbent

Alternatively, one might assume that the entrant can only access the low-cost portion of his cost curve over time, perhaps because the entrant needs time to ramp up his capacity or engage in learning by doing. Although they do not actually model that case, Williamson and Edlin express such a view.²⁹⁶ If so, then we do not have the case of an entrant who is initially just as efficient as the incumbent. Rather, the transfer of output to the entrant will be inefficient in the short run, and the case is actually one where entrant efficiency improves with passage of time rather than just output. In that sort of case, start-up costs have to be incurred to gain the human capital (experience) necessary to achieve the lower cost curve, and the effects will be as described in the previous two Subsections.²⁹⁷

D. *Ex Ante Effects on Incumbent Incentives*

The proposed restrictions on reactive above-cost price cuts do more than affect the likelihood of entry and the nature of post-entry competition. They also affect incumbent behavior pre-entry. Proponents have stressed the argument that the restriction's protection of entry by less efficient firms will force incumbent monopolists to lower everyday prices from a monopoly price to a limit price that is just low enough to keep out less efficient entrants.²⁹⁸ But while the restrictions may weakly encourage somewhat lower pre-entry prices, it is doubtful they will generally induce such incumbent limit pricing or that such a regime of enforced limit pricing is legally consistent with the argument for banning reactive above-cost price cuts. Further, proponents have ignored the other effect on pre-entry incumbent behavior—namely, that the restrictions reduce the incentives to create products that are so socially valuable that they make incumbents more efficient and earn them monopoly power. Such incentives will be reduced not only to the extent that the restrictions do induce lower pre-entry incumbent prices, but also because the restrictions will lower incumbent profits in the event of entry.

1. *The Likelihood and Legality of Encouraging Limit Pricing*

The proponent's conclusion that a restriction on reactive price cuts will lead to limit pricing (or a parallel increase in pre-entry output) depends on

efficiency in a way that does not suffice to overcome the incumbent's initial efficiency advantage. See *supra* Section IV.A.

296. See Edlin, *supra* note 6, at 975 & n.95, 977; Williamson, *supra* note 7, at 296, 298 n.43, 303-04, 313.

297. See *supra* Subsections IV.C.1-2.

298. See Edlin, *supra* note 6, at 946-47, 973-78; Williamson, *supra* note 7, at 308.

the premise that the restrictions will generally induce incumbent limit pricing that otherwise would not have occurred. This premise is dubious for several reasons.

First, not all prospective entrants will be less efficient. To the extent incumbents anticipate that some new entrants will be as efficient as (or more efficient than) the incumbent, those entrants are likely to enter no matter what above-cost pre-entry price the incumbent sets. This is also true for initially less efficient entrants who can raise the capital necessary to gain the experience to overcome their initial inefficiency. An above-cost limit price or output thus would sacrifice current monopoly profits without helping deter efficient entrants.²⁹⁹ Given the weak encouragement the restrictions provide to less efficient entrants, the lion's share of entrants will not be less efficient and thus incumbents will have little incentive to engage in pre-entry limit pricing.

Second, even if we restrict our attention to less efficient entrants, the restrictions are unlikely to increase significantly the likelihood that incumbents would adopt limit pricing. Incumbents contemplating limit pricing must calculate a trade-off between lowering their pre-entry profits and decreasing the risk that entry will lower their post-entry profits. Williamson assumed the latter would always govern but admitted that this was based on an "arbitrary assumption" that incumbents strictly prefer avoiding post-entry hazards to earning pre-entry profits.³⁰⁰ In fact, the preference is likely to run strongly in the other direction.

In part, this is because the pre-entry profits are earned in the present with certainty and thus should not have the time and uncertainty discounts a firm would rationally apply to any risk of a decline in post-entry profits.³⁰¹ Present value calculations can make the discounted value of any future loss of income from entry relatively small. Further, in a dynamic model, incumbents would not assume that today's cost and demand curves and entrant characteristics will prevail tomorrow. The market may be entirely changed by Schumpeterian competition, increases in entrant efficiency, decreases in barriers to entry, changes in consumer preferences, or sudden cost shifts. This uncertainty makes it rational to discount further any future profits that might be gained by deterring entry.

More important, though, is the low degree and magnitude of the *additional* risk of entry created by the restrictions on above-cost price

299. Those limit-pricing models that do conclude incumbents can keep out equally efficient entrants with above-cost prices, *see, e.g.*, CARLTON & PERLOFF, *supra* note 19, at 343-44 (summarizing the literature), use essentially the same model as Williamson, and thus fail for the reasons stated in Subsection IV.C.3, which rebuts the claim that a limit output or price will deter entry by a firm that shares the same declining cost curve.

300. Williamson, *supra* note 7, at 314.

301. *See* Areeda & Turner, *supra* note 56, at 1343-44.

cuts.³⁰² Incumbents will come to realize that less efficient entrants encouraged to enter by the restriction will be rare because their entry is futile in the long run. Further, incumbents will realize that if less efficient firms do enter, the incumbent (even with the restriction) can drive them out with a relatively minor time delay. It is highly unlikely it would be rational for the incumbent to sacrifice everyday high pre-entry profits to avoid this low additional risk of a brief interruption in those profits. That would require the incumbent to forgo permanently present certain monopoly profits on all its sales, in order to produce a small reduction in the uncertain risk that future entry will make the incumbent temporarily forgo a fraction of its sales. It would almost surely be more rational for the incumbent to fatten up on pre-entry monopoly profits, as such a strategy not only maximizes the incumbent's expected wealth but also assures enough reserves to deal with the wasteful losses from mothballing capacity that will occur when less efficient entry happens.

Pre-entry limit pricing would be even less attractive when the rational response to entry under a restriction will be not to try to compete with the less efficient entrant but rather to *raise* incumbent prices to hasten the time when prices can be reduced to drive out the entrant.³⁰³ In those cases, entry will not pose a risk of even temporarily lowered prices, though it will pose a risk of a temporary output decrease.

Still, while any encouragement to lower pre-entry incumbent prices will be weak, it does seem like restrictions protecting less efficient entrants at the margins may encourage created incumbents to charge lower pre-entry prices than they otherwise would have. This is because whatever calculation an incumbent makes in deciding whether to engage in limit pricing will include a somewhat larger likelihood of less efficient entry and larger costs when they do enter. Thus, sometimes lower pre-entry incumbent prices should result. Standing alone this will benefit consumer welfare. But, as the next Subsection shows, the restrictions also diminish the incentives to create incumbents with greater efficiency and lower costs, which will tend to increase pre-entry prices. Thus, the net effect on pre-entry consumer welfare is mixed.

There is also a legal oddity to the Edlin and Department of Justice position. As noted above, they argue that reactive above-cost pricing is predatory because it fits the *Grinnell* test of being designed to exclude rivals and maintain monopoly power.³⁰⁴ But that characterization would be equally true of the limit pricing they seek to induce incumbents to make.

302. To the extent that firms would engage in limit pricing with or without such a restriction, their limit pricing can hardly be claimed as a benefit of the restriction. It is only any *increased* likelihood of limiting pricing that matters.

303. See *supra* text accompanying notes 220, 238.

304. See *supra* Section I.B.

Setting a limit price has precisely the same effect on entrants and the same goal of maintaining monopoly power as a reactive price cut to the same price level.

Indeed, this led earlier courts that were attracted to the proposition that reactive above-cost price cuts could be predatory to the conclusion that limit pricing could also be predatory.³⁰⁵ This conclusion is surely misguided. We do not want antitrust courts in the business of forcing monopolists to raise their everyday prices. That would amount to a scheme of enforced monopoly pricing. But it does confirm that one cannot properly deduce whether pricing is predatory simply by determining whether it tends to keep rivals out of the market and maintain monopoly power.

The *Grinnell* test thus cannot itself support restricting reactive above-cost price cuts in order to enforce a regime of limit pricing. By the same token, the fact that current law permits limit pricing does not mean that limit pricing is affirmatively desirable or that we would want to force firms to adopt it. The lack of a legal ban merely means that trying to prohibit limit pricing would have undesirable consequences. In fact, affirmatively trying to require limit pricing would likely be undesirable, in part for the reasons discussed next.

2. *Reduced Incentives To Create Efficient Incumbents*

One must go one more *ex ante* step backward in time to consider the effects these restrictions would have on the incentives to create more efficient incumbents. The very premise that entrants are less efficient presupposes that this more efficient incumbent exists. But more efficient firms do not simply drop from the heavens. Someone had to make the risky investments necessary to create them in the first place. And their incentives to make those risky investments will be smaller if the law lowers the rewards for successfully creating a more efficient firm.

The proposed restrictions would lessen the rewards from creating a more efficient incumbent in numerous ways. First, to the extent proponents are correct that the enhanced threat of less efficient entry will induce incumbents to lower their everyday prices to keep out these entrants, then more efficient incumbents will reap lower everyday profits. Second, when a less efficient entrant does enter, the restrictions will prevent incumbents from adopting the above-cost prices that maximize their long-run profits. Where this has bite, it must lower the incumbent's expected profits and thus its rewards for having created a more efficient firm. Third, when faced with entrants that are equally efficient (or whose initial inefficiency will be overcome), the Edlin and Williamson restrictions will sometimes prevent

305. See, e.g., *Transamerica Computer Co. v. IBM*, 698 F.2d 1377, 1387 (9th Cir. 1983).

monopolists from offering the short-run profit-maximizing price post-entry, and thus will lower its returns. Fourth, when faced with efficient or inefficient entry, all the restrictions are likely to reduce the productive efficiency of the incumbent. This will also lower the incumbent's expected profits.

All this will lower the rewards for creating a more efficient incumbent. Faced with those lower returns, firms and investors will necessarily undertake less investment and innovation to try to create the next more efficient incumbent. Thus, the creation of more efficient firms will be reduced. This dynamic reduction in efforts to improve efficiency will lower productive efficiency and harm consumer welfare.

One might object that all this amounts to arguing that the law should not act to reduce monopoly profits. And so it does—when the monopoly profits are the fruit of having created a more efficient firm through desirable investment and innovation.³⁰⁶ We must remember that monopoly power is not itself undesirable. Market power simply means that the firm holding that power has a product so much more desirable or cheaper to provide than rival options that those other options do not constrain the firm from reducing output in order to raise prices and profits.³⁰⁷ And monopoly power just means a “substantial” or “significant” degree of market power,³⁰⁸ which merely means the firm has a product that is substantially more desirable or cheaper to provide than rival options. Creating a product that is substantially better or cheaper than rival options is highly desirable, since it leaves society far better off than it would have been had the product not been created.

Such monopoly power does not arise out of thin air. Someone had to invest or innovate under conditions of uncertainty to create a substantially better or cheaper product. And their incentives to take risks, invest, and innovate will be greater the larger their profits when they are successful. The ordinary rewards for doing so are the prospect of monopoly profits. We thus must be careful not to act as if the purpose of antitrust laws were to eliminate monopoly profits themselves. Such profits are an extremely valuable inducement to the creation of better or cheaper products.

This problem is particularly serious in high-technology markets, where such investments and innovation have the promise of not only creating something so valuable that it confers market power over preexisting rival

306. Of course, monopoly power can also be created in various anticompetitive ways, but if the antitrust laws are operating properly the incumbent monopolies should have achieved their monopolies through desirable means. And if the laws are not properly preventing the anticompetitive creation of monopoly power, then those laws are what need to be fixed.

307. See 1992 Horizontal Merger Guidelines, 57 Fed. Reg. 41,552 §§ 1.0-1.22 (Sept. 10, 1992); AREEDA & KAPLOW, *supra* note 132, at 556.

308. See *Reazin v. Blue Cross & Blue Shield of Kan., Inc.*, 899 F.2d 951, 967 (10th Cir. 1990); AREEDA & KAPLOW, *supra* note 132, at 448.

options, but may even generate a new market by creating a product much more desirable than preexisting market options. But the adverse effect on ex ante incentives is not limited to high-technology markets. It also exists whenever a firm has to decide whether to make investments in some old technology that will create a new facility that buyers will consider irreplaceable because of standard factors like transportation costs, or that will create market power because it satisfies a market niche that was previously unrecognized. Investments in changes in distributional methods or organizational form, personnel training, or the sheer creation of large-scale production methods can also lower costs or improve product quality in ways otherwise unattainable.³⁰⁹ In short, monopoly power can be desirably created in many low-tech and high-tech ways, and both of them will be discouraged if the ability to reap monopoly profits when successful is curtailed. This is true whether or not the innovation is patented, for the various restrictions on reactive price cuts would reduce monopoly profits on innovations whether or not they are manifested in patents.

To use the concrete illustration most important for predatory pricing purposes, consider the various market advantages that Edlin and the Departments describe incumbents as having in the airline industry: frequent daily flights, available connecting flights, economies of scale and scope, and brand-name advantages.³¹⁰ These are certainly advantages, but it is not as if they are undesirable or unearned. They rather reflect the desirable consequence of the incumbent making the necessary investments to produce a more valuable (or cheaper) product than its rivals.³¹¹ This is clearly true of developing a big enough network of flights to offer frequent and connecting flights and to take advantage of economies of scale and scope. It is even true for the market advantage that attends having created a more recognizable brand name, for any consumer willingness to pay more for a brand-name product indicates that the product is of higher quality as judged by consumers' revealed preferences.³¹²

To be sure, society would be even better off if it could have the more desirable or cheaper product and have it produced at cost. But that is a false

309. See SCHUMPETER, *supra* note 264, at 84-85, 88-89.

310. See *United States v. AMR Corp.*, 140 F. Supp. 2d 1141, 1149 (D. Kan. 2001); Memorandum of the United States, *AMR Corp.* (No. 99-1180-JTM), available at <http://www.usdoj.gov/atr/cases/f4800/4859.htm>; Edlin, *supra* note 6, at 943 n.12, 959.

311. More suspect are other advantages: for example, frequent flier programs and overrides paid to travel agents. *AMR Corp.*, 140 F. Supp. 2d at 1149. Both of these can be characterized as kickbacks that take advantage of agency problems to induce passengers to take less efficient flights. The frequent flier programs arguably induce individuals to spend more on business travel (the cost of which is billed to someone else or is shared with the government through tax deductions) in return for free personal travel. The travel overrides arguably reward travel agents with larger commissions for advising their clients to take more costly flights. But if either of these characterizations is true, then the proper remedy is not to ban above-cost price cuts but to ban the frequent flier programs and travel agent overrides that put passengers on higher-priced flights.

312. See *supra* Subsection IV.C.3.

choice. Unless given a high rate of return, firms will not invest to create the substantially more desirable or cheaper product. The monopoly power we are tempted to restrain will then never be created, but society will be worse off since it will be relegated to substantially worse or more costly market options. This problem with restrictions on reactive above-cost price cuts is really just a special case of the more general point that regulation (inside and outside antitrust) cannot take into account only the ex post effects that regulation has once a market and market power already exist. Regulation must also take into account any negative effect regulation has on ex ante incentives to invest and innovate to create something so valuable that it confers market power (over preexisting rival options) and may even generate a new market (by creating a product much more desirable than preexisting market options).

Limiting monopoly profits might seem desirable in a static model that focuses only on allocative efficiency. But in a dynamic model, such limits on monopoly returns will reduce productive efficiency, innovation and investment, and Schumpeterian competition to acquire temporary monopolies and the associated monopoly profits.³¹³ Moreover, here much of the reduction in monopoly profits does *not* result from improved allocative efficiency. Rather, the restrictions reduce monopoly returns in many ways that fail to confer such allocative efficiency.

The above concerns have tended to be missed by those advocating bans on above-cost predatory pricing because they adopt static assumptions about demand and cost curves and often seem to assume implicitly that the current incumbent is merely the undeserving beneficiary of those static market conditions. Indeed, as Baumol pointed out, limit pricing is generally only possible if an incumbent is a natural monopolist.³¹⁴ If a firm is truly a natural monopoly, antitrust law has little to contribute because it is impossible to create competition in such a market. Antitrust law can generally only contribute by protecting or restoring competition in markets that can support multiple firms, or by keeping free the even more important competition to create new product advantages that confer temporary monopoly power. Natural monopolies are by definition more durable. For them, the only real role of antitrust is to protect competition to become the natural monopolist. Such competition provides a market test that the monopoly really is natural, and that it remains so since today's natural monopoly can become tomorrow's temporary one if technology, costs, or demand changes sufficiently. Such competition also assures that the most efficient firm becomes the monopolist. But the hypothesis in these proposals is that the incumbent is as efficient as, or more efficient than, the

313. See generally SCHUMPETER, *supra* note 264, at 84-92, 99-106.

314. Baumol, *supra* note 7, at 11.

entrant (or else the entrant would hardly need protection), so such concerns are not at issue.

Instead, in natural monopolies inhabited by the most efficient firm, the grounds for regulating price are really no different than the traditional grounds for utility rate regulation. Not surprisingly, where for whatever reason the government has failed to institute such rate regulation, people who are otherwise attracted to such rate regulation might favor employing predatory pricing theory to try to fill in gaps in natural monopoly markets. But the most likely reason that rate regulation does not exist for any particular industry is that the government was not persuaded by the arguments for it. And if one thought such natural monopoly rate regulation were warranted, there would be no reason to limit it to cases in that industry where some claim of reactive price cuts provides the pretext. Moreover, triggering price regulation for all reactive price cuts risks applying it in cases that do not truly involve natural monopolies. It also means conducting such regulation through adversarial litigation and judges and juries who lack the ongoing involvement or expertise of utility regulators, or through other regulators who have not yet persuaded the legislature to give them the authority to engage in such general rate regulation.

One might imagine a different sort of objection that has not yet been made in writing by those proposing restrictions on reactive above-cost price cuts: Without a restriction, firms will engage in excessive investment and innovation. The argument could go as follows. With a restriction, those who create more desirable market options will still earn above-normal returns, but that return will be limited to the difference in efficiency between the newly created market option and other market options (for example, the next most efficient entrant). Without a restriction, monopolists will instead enjoy even higher prices, thus giving them a return greater than the value of the improvement over preexisting options that they created. This will give them incentives to make excessive investments in innovation, and thus dissipate the gains those investments confer.

The problem with this objection is twofold. First, innovations confer significant positive externalities that are not enjoyed by the innovator. Even innovators who get a patent only gain a right to exclude rivals for a limited number of years; after that, the social value of their improvement is completely appropriated by others. Further, innovations build on past innovations, meaning new innovations have a multiplier effect on future social benefits. Even during the patent period, patents can be invented around, or be faced with competing patents or independent innovations, so that no one who incurs the risk of investing in innovation is guaranteed a monopoly return. One could reply that this simply reflects the trade-offs (between rewarding innovation and disseminating its benefits) that the legislature made in defining patent law. But that argument cuts the other

way, for in making those trade-offs the legislature did not know that courts might later restrict patent holders who make above-cost price cuts, and thus the existing legislative trade-off assumes they get the full monopoly reward for their innovation. Moreover, many investments and innovations that create market power do not enjoy intellectual property rights at all. They may reflect improvements in methods of doing business that are nonpatentable or can be copied long before a patent period expires, and thus confer large positive externalities, the benefits of which will not be reflected in business decisions to invest. There is thus little reason to think firms would have excessive incentives to innovate unless their reward were limited to the difference between their costs (or product value) and the costs (or value) of the next-most-efficient firm. Rather, the process whereby each firm is rewarded with monopoly returns for making a product or production process that is better than preexisting options, and thus has incentives to engage in dynamic competition to replace each other over time, is more likely to be socially beneficial.³¹⁵

The second objection is that such excessive investments and innovation would be largely self-deterring. If the objection were true, it posits something like the following. Existing firms have costs of C . The innovator is thinking of making an investment that will give it costs of $C - I$, where I is the innovative improvement.³¹⁶ But after it drives out the existing firms, it will raise prices to M , thus enjoying a reward of $M - (C - I)$, or $I + M - C$, rather than just I . Thus, instead of making a (risk-adjusted and amortized) investment of up to I to create this innovation, it will invest more than I , up to $I + M - C$. But if the innovator did make an excessive investment E that cost more than I , then it could only recoup that investment with an expected price of $E + C - I$, which (since $E > I$) must be greater than C and thus greater than the cost of the existing firms and prevailing market prices. An investment that is expected to be unprofitable at prevailing market prices is unlikely to garner much capital funding. To anticipate that such excessive investments would be profitable, the firm would have to expect instead that it would initially price its product below a cost measure that included its investment costs in order to drive out the existing firms, and then raise prices to a higher level later. But such pricing could itself be challenged as below-cost predatory pricing, certainly under a total-cost approach and also under the approach laid out above, since those investment costs would be varied by the relevant alleged predatory increase

315. See *supra* Subsection IV.C.1.

316. One can make the same calculation for quality improvements. Assume that all firms have the same costs, that existing firms' product has a value $V = C$, and that the innovator invests in an innovation that offers a value $V + I$. But instead of pricing it at $V + I$, and thus enjoying a return of I , it prices it at M , enjoying a return of $M - V > I$. If it thus invests more than I in innovation, it will have to price at higher than $V + I$ to recoup that investment.

in output, which is everything the firm produces.³¹⁷ Further, a firm tempted to make such excessive investments in innovation must take into account the risk that other firms might innovate and improve their efficiency equally or more, thus restricting its anticipated monopoly returns. It will thus not have strong motives to invest in innovation with expected costs (of both the investment and postinnovation production) that exceed prevailing prices.

E. *Summary of Effects and Assessment of Possible Trade-Offs*

We can summarize the effects of a restriction on reactive above-cost price cuts in the following table:

317. See *supra* Part III. Since only some investments successfully create innovations, the measure would have to include risk-bearing costs, which should be reflected in the rates charged by capital markets.

TABLE 4. EFFECTS OF RESTRICTIONS ON REACTIVE
ABOVE-COST PRICE CUTS

	On Consumer Welfare	On Productive Efficiency
For Less Efficient Entry by Firms That Would Enter Anyway	Negative. Consumers of both incumbent and entrant pay higher post-entry prices.	Negative. Production shifted to less efficient entrant. Incumbent suffers uncompensated transition costs and decreased operating efficiency.
For Less Efficient Entry (Weakly) Encouraged by the Restriction	Mixed. Consumers who buy from entrant pay less in short run. Consumers who buy from incumbent may pay less or more.	Negative. Production shifted to less efficient entrant. Incumbent suffers uncompensated transition costs and decreased operating efficiency. Entry costs dissipated.
For Efficient Entry	Negative. Consumers of both incumbent and entrant pay higher short-term post-entry prices.	Negative. Decreased post-entry incumbent efficiency. Mixed but likely negative effect on incentives to create more efficient entrants.
For Entrants That Can Become Equally Efficient Only If Incumbent Efficiency Deteriorates	Mixed. Depends on whether increased allocative efficiency offset by increased costs.	Negative. Production shifted to less efficient entrant, and incumbent efficiency declines.
For Entrants That Can Become Equally Efficient Only by Increasing Own Efficiency	Negative. Consumers of both incumbent and entrant pay higher post-entry prices.	Negative. Production shifted to less efficient entrant in very short run, and incumbent suffers lower efficiency.
For Incumbent Pre-Entry Behavior	Mixed. May weakly encourage lower pre-entry prices, but also lessens incentives to create low-cost incumbents.	Negative. Decreased long-run incentives to create incumbents that are more efficient than pre-existing market options.

These conclusions differ from those of traditional analyses, which generally have assumed instead that restrictions on above-cost price cuts exchange certain short-term post-entry costs for an uncertain long-term post-entry gain.³¹⁸ Instead, the restrictions confer no long-term post-entry gain and can inflict long-term costs, and while they may sometimes confer a short-term post-entry gain, on other—and more frequent—occasions they inflict a short-term post-entry cost. Further, the restrictions may lower or raise pre-entry prices, but also impose serious pre-entry costs by reducing the creation of more efficient incumbents and entrants.

While these effects cannot logically exclude the possibility that the restrictions may have net desirable effects in some cases, they do suggest that it is extremely unlikely that the overall results of the restriction would be desirable. First, the possibly beneficial effects on pricing are mixed, whereas most of the negative effects are unambiguous. Second, where these mixed effects do prove on balance beneficial, the effects are weak because the restrictions encourage little additional entry by less efficient firms. In contrast, the lion's share of firms will be efficient, and for them the effects on pricing are unambiguously negative. Third, the possible benefit to post-entry pricing in the case of the encouraged less efficient entrants is short-run, since they cannot survive in the long run once the restriction on above-cost price cuts expires. In contrast, the harms to pricing that result from discouraging the creation of more efficient entrants and incumbents are long-term.

Fourth, even if there were a net benefit to consumer pricing, it would have to be weighed against the clear loss to productive efficiency. For reasons discussed above, trading off increased consumer welfare for lowered productive efficiency is generally not desirable even if we assume a straight trade-off between productive and allocative efficiency over a similar time frame. It is even less likely to be desirable when the consumer welfare benefits are mixed, weak, and short-run, and apply in a limited set of cases compared to a clear loss in productive efficiency over a longer period that covers a broader set of cases. Fifth, the loss of productive efficiency is not merely static but dynamic, undermining a competitive process of innovation whereby each firm has incentives to lower costs further or improve product performance, a process that confers enormous positive externalities on society.

Indeed, Richard Schmalensee showed some time ago that “privately profitable entry may not be socially desirable if the entrant's costs exceed those of existing firms” because it can worsen productive efficiency more

318. See, e.g., *Barry Wright Corp. v. ITT Grinnell Corp.*, 724 F.2d 227, 234 (1st Cir. 1983) (Breyer, J.) (analogizing the restrictions to sacrificing a bird in the hand for two in the bush); Harold Demsetz, *Barriers to Entry*, 72 AM. ECON. REV. 47, 56 (1982) (same); sources cited *supra* note 60.

than it improves consumer surplus.³¹⁹ He concluded that it could well be that “[s]ociety as a whole would be better off if existing firms were allowed to bribe potential [higher cost] entrants not to enter, or if entry were restricted by government regulation of some sort.”³²⁰ And this was under the assumptions that such entry was permanent, expended no entry costs, and inflicted no transition costs or efficiency losses on incumbents, and free post-entry price competition was allowed. Where instead the less efficient entry is induced by a temporary price umbrella that worsens short-term price effects, makes long-term benefits futile, dissipates entry costs, and may reduce incumbent efficiency or discourage its creation, there is even more reason for skepticism. And the trade-off is even more likely to be negative when one also considers the effects on efficient entrants.

Finally, to the extent there are beneficial effects, they basically boil down either to the “infant firm” argument that new firms need to be encouraged because capital markets underrate them, or the case of a natural monopolist who cannot be threatened by an entrant who is equally or more efficient, and is thus a good candidate for utility rate regulation. But the proposed restrictions are poorly tailored to advance those goals. They apply even to industries where entrants need no encouragement, and fail to protect new firms that do not face incumbent monopolists who make reactive price cuts. And they extend well beyond natural monopoly cases, do not cover all the natural monopoly cases one might wish to regulate under such a theory, and are less likely to induce the correct rate. Thus, these benefits can more readily and accurately be achieved either through general rules to protect or subsidize new firms or through utility rate regulation. Where such regulation already exists, the proposed restrictions will not have these possible benefits. Where it does not exist, it would seem to reflect a societal judgment that protecting entrants or regulating rates is unwise—a judgment we have no warrant for overturning through antitrust law.

F. *The Restrictions Cannot Reasonably Be Construed or Modified To Eliminate or Suspend the Market-Power Requirement*

As the above analysis indicates, one key problem with the restrictions is that, because monopoly or market power is required before a firm's prices can be regulated under any antitrust or competition law, any restriction on above-cost predatory pricing that hopes to protect less efficient entrants must be futile in the long run. One might thus be tempted to dispense with the monopoly or market-power requirement. But this would require a

319. See Richard Schmalensee, *Is More Competition Necessarily Good?*, 4 INDUS. ORG. REV. 120, 120 (1976).

320. *Id.*

statutory or treaty amendment. Further, as we shall see, this requirement is not merely an artifact of the particular proposals made. It is rather a necessary feature of any doctrine of predatory pricing. Nor would any such modification eliminate all the adverse effects of a restriction on above-cost incumbent pricing.

To begin with, monopoly or market power is required by existing competition law, which does not restrict even below-cost predatory pricing unless the actor has monopoly or market power. To be sure, we could change that law. But we could not do so through case law. It would require a statutory amendment in the United States or a treaty amendment in Europe. Under U.S. law, a claim that unilateral pricing decisions constitute monopolization (or attempted monopolization) under section 2 of the Sherman Act requires proof not just of predatory pricing but of monopoly power (or enough market power to create a dangerous probability of acquiring monopoly power).³²¹ Likewise, European law requires proof of a dominant position to make predatory pricing actionable under Article 86.³²²

More important, any amendment eliminating the market-power requirement would be unwise. Without such a requirement, a doctrine of predatory pricing would effectively aim to regulate all reactive pricing on competitive markets. Such competitive pricing is precisely what the antitrust laws seek to foster on the grounds that competitive markets can best set prices. Competitive firms are supposed to compete by each trying

321. See, e.g., *Brooke Group Ltd. v. Brown & Williamson Tobacco Corp.*, 509 U.S. 209, 222 (1993) (defining predatory pricing in terms of monopolization); *Spectrum Sports, Inc. v. McQuillan*, 506 U.S. 447, 455-56 (1993) (defining attempted monopolization); *United States v. Grinnell Corp.*, 384 U.S. 563, 570-71 (1966) (defining monopolization). Some have argued that the recoupment requirement itself seems to impose a higher market-power requirement, and may thus eliminate or constrict any claim of attempted monopolization through predatory pricing. See 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 724a-b, at 284-85, ¶ 728b. Monopoly power is not necessary under the U.S. Robinson-Patman Act, but even it requires some level of market power. *Brooke*, 509 U.S. at 222 (stating that the defendant must have enough market power that its predatory pricing creates at least "a reasonable possibility of substantial injury to competition"). The recoupment requirement may elevate the market-power requirement even further. See 3 AREEDA & HOVENKAMP, *supra* note 15, ¶ 724a-b, at 284-85, ¶ 728b. Moreover, the Robinson-Patman Act has various statutory limitations that make it a poor vehicle for generally policing predatory pricing. In particular, the Robinson-Patman Act is limited to price discrimination (and thus does not cover a uniformly predatory price) and commodities (and thus would not cover airline transportation or other services). 15 U.S.C. § 13(a) (2000). Further amendments would be necessary to restrict above-cost predatory pricing. In particular, the Robinson-Patman Act specifically allows price cuts to match competition in good faith, *id.* § 13(b), which directly contradicts the core of these proposals to restrict reactive above-cost pricing. The Act also allows different prices based on varying costs or market conditions. *Id.* § 13(a).

322. See *Case C-62/86, AKZO Chemie BV v. Commission*, 1991 E.C.R. I-3359 ¶¶ 34-74 (E.C.J.). That decision held that a market share over fifty percent would suffice. *Id.* ¶ 60. Market shares below fifty percent might also constitute a dominant position depending on other structural factors that affect the degree to which market shares imply market power. See *Joined Cases C-68/94 & C-30/95, French Republic v. Commission*, 1998 E.C.R. I-1375 ¶¶ 111, 242-248 (E.C.J.); *Case T-102/96, Gencor Ltd. v. Commission*, 1999 E.C.R. II-753 ¶¶ 202, 239-263 (Ct. First Instance); IVO VAN BAEL & JEAN-FRANÇOIS BELLIS, *COMPETITION LAW OF THE EUROPEAN COMMUNITY* ¶¶ 248-252 (3d ed. 1994).

to match and then beat the price and quality of their rivals. To interfere with this dynamic is to interfere with the "central nervous system of the economy,"³²³ and to "set sail on a sea of doubt" by requiring courts to determine what reasonable prices are on competitive markets.³²⁴ Moreover, once one dispenses with the market-power requirement, it is not clear which firm in any given market would be subject to the restriction on reactive above-cost predatory pricing, since all the firms in the competitive market are reacting to each other. One could try to ban *all* firms from reacting to new entrants, but then one faces the question why the law should so favor entrants (which would seem to lead to inefficient overinvestment in entry into competitive markets), and whether any new entrant would not then become immediately an incumbent forbidden from engaging in reactive price cuts. The result would be to ossify and distort pricing on competitive markets. Further, while such a rule would preserve the long-term existence of the entrant, all the other ill effects from imposing post-entry price floors would continue to apply.

Another possible modification would concede that any price restriction must expire when the monopoly power erodes, but provide that once the incumbent's postexpiration above-cost price cut causes the incumbent's market share to grow back over the monopoly threshold, that above-cost price would amount to attempted monopolization. One might hope through this sort of regulation to keep the incumbent perpetually shy of a monopoly share. But this modified approach would raise new problems because the illegal pricing decision would be neither a price cut nor reactive to entry. Since the incumbent's postexpiration price cut would initially be legal, the law would have to make illegal the incumbent's failure to impose a price *increase* (or output decrease) once the incumbent got back to a market share close to monopoly power. This hardly seems likely to promote consumer welfare. Other problems would result because the rule would no longer be triggered by a reaction to entry. The moment that retriggers price regulation would become obscure, with the modified rule putting the incumbent at great peril for not guessing accurately when a court or jury would deem it on the verge of crossing the line into monopoly power again. Nor would the right baseline for a legal price or output be clear, since it would no longer be the price or output that just preceded the moment of illegality. And again, while any benefits from entry would no longer be merely short-run, all the adverse effects of the restrictions would continue to apply.

One might be tempted to respond to this problem with an amendment providing that, while the bans on above-cost predatory pricing apply only to

323. *United States v. Socony-Vacuum Oil Co.*, 310 U.S. 150, 226 n.59 (1940).

324. *United States v. Addyston Pipe & Steel Co.*, 85 F. 271, 284 (6th Cir. 1898), *aff'd*, 175 U.S. 211 (1899).

incumbents who *begin* with monopoly or market power, those bans continue to restrict those incumbents even after they lose their monopoly or market power. However, the effects of such a modified proposal would be even worse. The short-run effects would be the same as the existing proposals. But the long-run effects would differ. If never able to cut prices to match or beat the entrant, the incumbents would necessarily be driven from the market in a way that will not permit reentry. The market will thus be left to the new entrant who is, by hypothesis, less efficient. That new entrant will have incentives to raise prices to its own monopoly level, which will be higher than pre-entry incumbent prices because its marginal costs are higher. Thus, in the long run, rather than just being futile (like the existing proposals), the modified proposal would affirmatively harm productive efficiency and consumer welfare. In the even longer run, under this modified proposal, other less efficient entrants might enter and produce a competitive market full of less efficient firms. If so, a low-cost monopoly would be replaced by a high-cost unconcentrated market, which is unlikely to be desirable for reasons discussed above.³²⁵

Alternatively, one might try an amendment providing that the restriction on above-cost predatory pricing applies only to incumbents who begin with monopoly and market power, and persists even after they lose that power, but lasts only for some fixed period of time, like the twelve to eighteen months suggested by Edlin. But there are two possibilities under such a proposal. The first is that being forced to mothball its capacity for this time has reduced the incumbent's efficiency to the point that it no longer has an efficiency advantage over the entrant. If so, then such a modified proposal would have the same effect as a permanent restriction on above-cost price cuts. The second is that this enforced mothballing of incumbent capacity has not eliminated its efficiency advantage. If so, then the regulation would again be futile because at the expiration of twelve to eighteen months (or whatever period is specified) the incumbent would again lower prices and drive out the entrant. The restriction will still have unambiguously negative effects on productive efficiency and inflict a wasteful contraction of production that will just have to be reexpanded. The effects on allocative efficiency will also remain unambiguously adverse for any entrants who would have entered even without the restriction, and the encouragement for additional entry by less efficient firms will remain weak. It would be the case, however, that when a less efficient entrant *is* encouraged to enter, the effects on allocative efficiency would be positive, since such a modified proposal at least would not give the incumbent affirmative incentives to raise prices in order to speed the day when the restriction expires.

325. See *supra* Subsection IV.C.1.

V. UNAVOIDABLE IMPLEMENTATION DIFFICULTIES WORSEN THE ABOVE EFFECTS

In addition to the above fundamental problems posed by any restriction on reactive above-cost prices, the restrictions also raise many implementation difficulties that worsen their likely effects. But this is not because the restrictions are poorly formulated. It is because any formulation must make choices about how to deal with unavoidable problems that will bedevil any effort to regulate above-cost predatory pricing. No matter what system is adopted, it would somehow have to ascertain the moment of entry that triggers the restriction, deal with quality changes designed to evade it, and define a post-entry price floor (or output ceiling) that will lead to inefficiencies unless precisely adjusted for changing market conditions. In doing so, there are no perfect choices. Rather, any choice will inflict some significant distorting effect on entry, innovation, and efficient pricing. One can try to adjust the restrictions to minimize these additional inefficiencies, but one cannot eliminate them, and they will worsen the adverse effects of the proposed restrictions. These additional inefficiencies are sufficiently large that they would offset any gains from such restrictions even if, contrary to the above analysis, such net gains likely existed.

Past attention has focused on administrative problems in defining the post-entry price floors and output ceilings. But even bigger problems result from difficulties in defining the moment of entry and controlling for possible quality distortions. If the moment of entry that triggers the restriction is defined to be when the entrant actually begins sales (or at any other time after entry is foreseeable), the incumbent will simply be able to cut prices beforehand, rendering the restriction ineffectual and even less likely to encourage entry. If the moment of entry that triggers the restriction is defined to be when entry is first foreseeable, then the law would be raising incumbent prices during a period where this is not offset by lower entrant prices, thus worsening the likely mix of effects. Further, if the moment of entry is given such an early definition, then either a twelve- to eighteen-month period of restriction would often expire by the time the entry starts (making the restriction ineffectual), or the period during which incumbent prices are restricted will have to be lengthened, worsening the inefficiencies that result from creating price or output inflexibility in the face of changing market conditions. Another huge problem is that any price or output floor will provoke inefficient increases in product quality, and any effort to clamp down on that by restricting product enhancements will hamper productive efficiency.

Further, the administrative problems with defining the price floors and output ceiling are underappreciated in two ways. First, it is not merely a matter of judgment whether the administrative problems with any flexible

price floor or output ceiling outweigh the administrative problems with a cost-based rule—rather, any flexible floor or ceiling necessarily creates larger administrative costs because implementing it requires assessing price-output possibilities up and down supply and demand curves that change over time rather than comparing one price to costs at a given output level. Second, these administrative problems cannot be avoided by tweaking the proposals. They are rather an inherent consequence of trying to regulate incumbent pricing or output. Whether the law uses a flexible standard or rule, the regulated price or output will often fail to reflect changing market conditions accurately and will thus produce additional inefficiencies.

A. *When Is the Moment of Entry?*

Under all the approaches for restricting reactive above-cost price cuts (or output increases), the restrictions are triggered by entry. But the moment of entry is not so easy to define. Is it when the entrant first announces its entry? When it first applies for a permit or license? When it begins construction on a new plant? When it begins its marketing campaign? When it sells its first test product? Or when it first attempts a substantial quantity of sales? Edlin is the only proponent to address this definitional difficulty, and he takes varying positions on it. In analyzing one case, he states that the attempt to enter did not qualify because the entrant never got to the point where it actually produced the product.³²⁶ In another case, he concludes that beginning construction suffices to trigger the ban even though the entrant had not yet sold the product.³²⁷ Either position raises problems, which are only exacerbated by ad hoc shifts from one position to the other.

Suppose one picks one of the later moments as the true moment of entry. Then the problem is that at one of the earlier moments the incumbent will know entry is forthcoming and thus can lower prices (or expand output) in anticipation. The restriction on reactive price cuts will be toothless because the incumbent can react before the defined moment triggers the restriction. For example, if the entry is defined by actual production, then the incumbent can just wait until construction is almost completed and cut prices before the entrant ever sells anything. If so, the restriction becomes ineffectual and is even less likely to encourage entry or have any beneficial effect.

To deal with this problem, Edlin effectively creates an ad hoc rule. In one case, cutting prices before the entrant makes any sales is inappropriate

326. Edlin, *supra* note 6, at 987-88.

327. *Id.* at 988.

because the entrant's construction plans made it "substantial."³²⁸ In another case, cutting prices before the entrant makes any sales is acceptable because one can infer the entrant was "insubstantial" from the fact that a buyer with fifty-percent market share accepted a five- to ten-percent price cut from the incumbent.³²⁹ Edlin bases the latter conclusion on the assumption that such a big buyer's incentives are largely aligned with consumer welfare.³³⁰ But this inference of efficiency is probably untrue because powerful buyers often have incentives to cut deals that benefit themselves even though they create seller market power.³³¹ In any event, this approach introduces additional sources of great uncertainty. Just which buyers are large enough that their acceptance of a reactive price cut justifies deeming entrants "insubstantial," and what are the other situations where an inference of efficiency will justify suspending the ban on reactive price cuts?

Suppose one instead picks one of the earlier moments of entry, such as announcing entry or applying for a permit or license. This raises many other problems. First, with such an early definition of the moment of entry, the restriction will impose an incumbent price floor, with all the adverse effects on pricing and efficiency, well before the entrant actually makes any sales. This reduces further any likelihood that the benefits resulting from encouraging entry will outweigh the adverse effects, for the period of possible benefit will be shorter than the period of adverse effects.

Second, such an early definition of the moment of entry may also make the restriction ineffectual. In particular, any definition that tracks an entry announcement, application, or even construction will often mean that the moment of entry occurs more than eighteen months before the entrant actually seriously sells its product. But the Edlin and Williamson restrictions only last twelve to eighteen months at the outside. Thus, with such an early definition of the moment of entry, these restrictions would be likely to expire before the entrant ever seriously sells its product, and thus cannot prevent an incumbent from adopting a reactive price cut after the entrant starts selling.

The profit-maximizing price floors do not raise this problem since they set no expiration time. But they produce a different anomaly. The incumbent's prices would have to be monitored for a long period of time before actual entrant sales commenced in order to make sure the incumbent came sufficiently close to maximizing short-term profits. Such monitoring is costly. Moreover, since the entrant would not yet be making sales, the

328. *Id.*

329. *Id.* at 987-88.

330. *Id.*

331. See, e.g., 4 AREEDA ET AL., *supra* note 218, ¶ 943b, at 204-06 & n.4; Herbert Hovenkamp, *Mergers and Buyers*, 77 VA. L. REV. 1369, 1375-76 (1991). This is just an application of the Coase Theorem. A powerful buyer and seller will have incentives to make an agreement that preserves supracompetitive pricing and divides the profits among them.

price that maximizes short-term profits would be the monopoly price. Thus, such a restriction would mean that for a substantial period the government would be mandating monopoly pricing. Even if we want to encourage entry, it is hard to believe we want to do so by giving potential entrants an entitlement to require incumbents to charge monopoly prices before the entrant makes any sales. Further, lengthening the period of restriction will only worsen all the administrative problems of applying such price floors in the face of changing market conditions.³³²

Third, any early definition of the moment of entry will make the restrictions vulnerable to strategic exploitation. By merely announcing entry or making an application, any firm can restrict the prices of another firm (and under the Edlin proposal can freeze their prices and quality too). If one tries to avoid this by restricting the moment of entry to credible announcements or committed applications, then one has the ambiguity of just which announcements or applications are credible or committed enough to trigger the restriction, and just how incumbent firms are supposed to predict what antitrust litigation will in the future conclude on that topic. Picking some middle moment like actual construction of a new facility might work for some markets, but even when it does, it lends itself to reactive price cuts after the announcement or application but right before construction begins. And just when construction begins might itself be ambiguous.

Moreover, even if the prospect of future entry has been made certain by the announcement or application, how can an incumbent know whether the coming entrant will actually offer the twenty-percent discount necessary to trigger the Edlin rule? This seems especially uncertain since, under Edlin's own analysis, differences in quality might make a nominal twenty-percent price difference insufficient.³³³ Even if the entrant says it will offer a twenty-percent price discount and the same quality, such announcements are unreliable, nonbinding, and may be made purely strategically to freeze their rivals. Here Edlin creates another ad hoc exception. Although no twenty-percent price discount has been offered, the "substantial" entry requirement should be deemed satisfied if the entrant has construction plans to serve most of the market, with the price freeze lifted if the entrant turns out not to sell at a twenty-percent discount.³³⁴ This allows entrants to freeze rival prices by mere construction even though they have not undercut incumbent prices at all, and the creation of another ad hoc exception again undermines any certainty the rule might have had.

332. *See infra* Section V.C.

333. *See infra* Section V.C.

334. Edlin, *supra* note 6, at 988.

All these problems are multiplied if one triggers the restrictions not only in cases of actual entry but, as Williamson would, in cases where a “fringe firm[]” makes a “new investment” significant enough to be considered tantamount to entry.³³⁵ The impulse is understandable because the economic effects of such investments and entry may be the same. But it exacerbates uncertainty when incumbents cannot be sure which rival investments will be considered significant enough to trigger above-cost restrictions, and it widens opportunities for strategic gaming when announcing any new investment might freeze the output of a dominant firm.

The problem is not an avoidable one. To make them plausible, all the proposals have to start with some moment of entry to trigger the restriction on reactive price cuts. Otherwise, they would amount to a general regulation of pricing that is entirely inconsistent with a market approach. But no matter which moment one picks, the restriction either becomes toothless (eliminating any benefits) or lengthens the period of adverse effects and produces strategic behavior and anomalous results. If the moment of entry is defined to occur either when actual sales are made, or earlier when entry is planned with a short period of restriction, then the rule cannot really prevent the incumbent from adopting reactive price cuts. If the moment of entry is defined to occur earlier than when sales are made, then the period during which consumer prices are elevated by the price floor will exceed the period during which the entrant might lower prices, and anomalies and strategic abuses become possible. And any early definition that lengthens the period of restriction would worsen the difficulties in defining the incumbent price floor or output ceiling.

B. *Post-Entry Quality Changes*

Whenever prices are regulated, firms predictably shift to nonprice competition. For example, back when airline prices were thoroughly regulated, airlines competed with fancy meals and more frequent, less crowded flights. More generally, one can expect firms whose prices are regulated at above-cost levels to compete by improving the quality of their product.³³⁶ This complies with the restrictions on price cuts but effectively lowers the quality-adjusted price in a way that still allows the incumbent to drive out the less efficient entrant. But because the restriction prevents price cuts that otherwise would occur, it inevitably induces the creation of products that make a different quality-price trade-off than buyers would prefer on a free market, and these quality improvements are thus inefficient.

335. Williamson, *supra* note 7, at 292 n.26.

336. They might also increase advertising or service levels in a way that is equivalent to a quality increase.

The U.S. Departments' proposals were partially responsive to this problem, regulating not just prices but airline capacity. Airlines would thus not be able to respond to an entrant by just adding planes to provide a more convenient schedule, which is one way of improving quality. But airlines could still have evaded this restriction by offering more frequent flights on smaller planes, which would be inefficient but still offer fliers more flexibility while complying with the capacity limit. Or airlines could inefficiently improve quality in other ways, with fancier meals or service. All these quality improvements would be inefficient because (to the extent the restriction on price cuts has bite) they would be replacing a price cut that consumers would prefer to the quality enhancement.

Thus, the U.S. Departments' approach has the problem that generally bedevils efforts to restrict nonprice competition. Whenever one tries to clamp down on one form of nonprice competition, the underlying incentives drive firms to whatever forms remain unregulated. For example, when airline regulators tried to make their price regulations meaningful by clamping down on nonprice competition, they specified that airlines could only offer "sandwiches" on international economy flights. Airlines responded with such tactics as putting duck à l'orange on one slice of bread for an open-faced "sandwich." The Williamson output-ceiling faces similar problems. Firms will have incentives to evade the output ceiling (and the effective price floor that it implies) by increasing quality.

A firm might even have incentives to change its product so much that it can argue it has a new product that is not subject to the price or output restriction. This can create incentives to inefficiently improve or even worsen the product. If courts respond by subjecting new products to the restrictions imposed on old, related ones, then the rule will deter genuine innovation.

Edlin attempts to address this problem by banning incumbents not only from cutting prices but also from making any "significant product enhancements."³³⁷ But this creates severe administrability problems. Just how is the antitrust court or jury supposed to decide which product enhancements are "significant," or more to the point, how is the incumbent supposed to be able to predict what a future unknown judge or jury will later decide was "significant"? Further, what is a court supposed to do if the incumbent says it is not enhancing the old product but introducing a new one?

Moreover, to the extent this restriction on product enhancements is administrable, it is undesirable. It achieves the aim of lessening nonprice competition that might undermine a price floor at the cost of lessening all product innovation. Even if the price floor did seem well designed to

337. Edlin, *supra* note 6, at 945.

benefit consumer welfare by restraining monopoly pricing, that rarely is as important as the efficiency benefits of product innovation.³³⁸ But where, as here, the price floor's contribution to consumer welfare is probably harmful (or at least questionable), there is no sensible reason to sacrifice the productivity gains of innovation in order to maintain the price floor.

Indeed, under the Edlin approach, a less efficient firm that realizes a wave of innovation is forthcoming has incentives to enter in order to trigger the prohibition on incumbent product enhancements. This not only has undesirable effects on incumbent innovation, it also encourages costly and inefficient entry by a firm that would not enter but for the ability to freeze the innovation of others. This is undesirable enough when the innovation affects only one market. Because innovation in one market in fact often ends up having applications to other markets and sometimes even redefines the markets, it raises the even graver concern that a firm in a related market might enter the incumbent's market to freeze innovation that might pose a competitive threat in that related market. This worsens the likely effects of any encouraged entry.

Again, the problem is an unavoidable one. One can leave quality competition largely unconstrained, which makes the price or output regulation ineffective at achieving its goal of encouraging entrants but harms customers by depriving them of the lower price-quality trade-offs they prefer. Or, one can really clamp down on quality competition, which makes the regulation more effective, but at the excessive cost of eliminating product innovation.

C. *Difficulties in Defining the Incumbent Price Floor or Output Ceiling*

The approaches that set the incumbent's post-entry price floor in relation to the price that would maximize short-term profits raise plain administrability problems. Determining which price maximizes profits is highly uncertain and variable over time. True, critics of cost-based tests are correct that judging incremental costs can also be administratively difficult. But determining the profit-maximizing price requires determining not just the costs that were incurred at the marginal output level, but the costs all along the supply curve at every possible output level. Thus, such profit-maximizing price floors multiply all the complex problems about projecting costs, distinguishing between fixed and incremental costs, allocating common costs, and evaluating capital costs and risk. Worse, determining the profit-maximizing price also requires ascertaining the incumbent's demand curve at each price and output point, as well as the extent of

338. See *supra* Sections IV.C-D.

incumbent market power, to determine just what price equilibrates marginal revenue and cost. And determining the incumbent's demand curve necessitates knowledge not just of total buyer demand at each price (which will turn not only on their inherent preferences, but their willingness to switch to substitute products or markets as well), but also of what outputs and prices rivals would offer at each incumbent price. Further, each of those curves, and thus the profit-maximizing price, will change from day to day as market conditions or technologies change.

In short, it is not merely a matter of judgment whether the administrative problems with any flexible price floor outweigh the administrative problems with a cost-based rule, even if one does not agree with my effort in Part II to clarify cost measurements. Any flexible floor must take into account changing market conditions and consider price-output possibilities up and down the changing demand and cost curves. A cost-based rule need only compare, at one actual output point, the incumbent's price to its actual costs.

Given these difficulties, there is probably no practical way to determine any difference between an above-cost price and the short-term profit-maximizing price.³³⁹ Firms have trouble enough in making such judgments, but they are in the business of doing so and in the end are policed by markets that weed out the firms that tend to guess wrong. Regulators are not. Worse, if made an antitrust claim, the issue will be left to antitrust courts that will have even greater difficulty since they are not (like a regulator might be) a single entity with the expertise and power to continuously monitor and prospectively approve price levels. Instead, antitrust courts will be regulating prices through the clumsy vehicle of adversarial lawsuits that involve varying judges and juries asked to retroactively assess claims that a particular set of prices was too low. Such a cumbersome litigation process would be highly burdensome on courts and impose direct costs that firms would pass on to customers. It would also cause uncertainty that, to avoid the risk of treble damages, will incline incumbents to charge higher prices than they otherwise would have, thus harming consumer welfare.³⁴⁰

The U.S. Departments tried to avoid these problems by banning price cuts only if they are "clearly" or "substantially" lower than the short-term profit-maximizing price.³⁴¹ This should make incumbents less risk-averse

339. See *Barry Wright Corp. v. ITT Grinnell Corp.*, 724 F.2d 227, 234-35 (1st Cir. 1983) (Breyer, J.); 3 ARFEDA & HOVENKAMP, *supra* note 15, ¶ 736c2, at 381-83; Joskow & Klevorick, *supra* note 71, at 255.

340. Antitrust courts have consistently rejected any legal theory that requires them to monitor the day-to-day reasonableness of prices under changing market conditions as inadministrable for courts, burdensome on litigation, and too uncertain for business planning. See, e.g., *United States v. Trenton Potteries Co.*, 273 U.S. 392, 397-98 (1927).

341. See *supra* Section I.A.

about pricing down to their short-term profit-maximizing level. But around any price floor there will be an inevitable zone of uncertainty. And here the zone is great because it depends not only on just what adjudication might conclude about the short-term profit-maximizing level, but also on the vague terms “clearly” or “substantially,” which will likely mean something different to every adjudicator or juror who applies them. This approach does not eliminate the ambiguity; it just moves the ambiguity to a different price point, and worsens the degree of ambiguity to boot.

The U.S. Department of Justice also emphasized that American Airlines was forgoing an option it had itself decided was more profitable in the short run.³⁴² But to the extent the rule hinges on the availability of such internal documents, all it will do is drive profit calculations underground, thus leaving the rule ineffectual. If it does not hinge on the existence of internal documents, then the rule will remain inadministrable and uncertain, deterring incumbents from making desirable price cuts. These effects will be particularly undesirable in the cases of entry that really matter for the long run: when the entrant is (or soon will be) just as efficient as the incumbent.

The European cases may likewise be trying to escape these problems with their emphasis on the selectivity of the price cuts.³⁴³ But the rationale for this possible limitation is unclear. If a selective price cut really does not alter prices elsewhere, it must be because the selected area is its own market. The price cut is then simply occurring in the market where entry occurred, which is not much of a limitation. Perhaps the European authorities have in mind that the selectivity of the price cut means that the uncut prices in other areas provide an objective benchmark as to what price level does maximize short-run profits. However, while such selectivity (as those authorities at points suggested) helps dismiss the possibility that the price cut was prompted by some cost reduction rather than by entry, the fact is that the short-term profit-maximizing price in a market with an entrant will differ from that price in other areas where there is no competitor. So the selectivity of price cuts cannot avoid the administrability problem of determining what the short-term profit-maximizing level is. And it adds the administrability problem of determining just when pricing is sufficiently “selective” to invoke the rule.

Edlin tries to get around these well-known problems with a flat rule. The incumbent cannot charge any price lower than its pre-entry price if the entrant has offered a 20% discount. But this also has serious problems. Even if the nominal price is clear, the effective price will vary with differences in service, credit, or delivery associated with the product. There

342. *United States v. AMR Corp.*, 140 F. Supp. 2d 1141, 1152-53, 1155, 1181 (D. Kan. 2001).

343. *See supra* Section I.A.

will also be ambiguities about the nominal price whenever the incumbent varies prices or sells a variegated product. For example, in the airline industry that provoked these proposals, a seat on a plane is sold at widely disparate rates depending on purchaser identity, advance purchase, Saturday stayovers, restrictions on changes, and the competing demand from customers in all the other cities that fly through that route. The last factor means that the price for a seat from the hub city to spoke city *A* turns not just on demand for travel between the hub and spoke *A*, but also on demand for travel between spoke cities *B-Z* and spoke city *A*.³⁴⁴ There is no one single price to pick. If courts tried to pick an average price, they would raise prices for the whole array (roughly half) of customers who otherwise would have paid a lower price. Similar problems would apply if the incumbent varied financing or credit terms for different buyers.

These problems are multiplied by the need to compare the pre-entry incumbent schedule of prices to entrant prices to determine whether the entrant prices are 20% lower. Indeed, the need for that comparison introduces a new problem: The entrant product might be of lower quality. Edlin recognizes that this will require a quality adjustment to determine whether the entrant has offered a “20% quality-adjusted discount.”³⁴⁵ These quality adjustments are significant enough that a 25%-40% price difference only “probably” qualifies as a 20% discount.³⁴⁶ But once one introduces this vague assessment of quality adjustments, any supposed administrative simplicity vanishes. The problem is even worse if one resorts to Edlin’s alternative standard that the entry has offered a “substantial” discount,³⁴⁷ a vague placeholder whose definition can vary widely from tribunal to tribunal.

More importantly, to the extent the pre-entry price and 20% discount trigger can be established, setting a price floor for the incumbent (and price ceiling for the entrant) has obvious inefficiencies. Prices in all markets vary with rapidly changing costs, technologies, and demand. Requiring firms to stick to price floors and ceilings thus rapidly produces inefficiency. For example, if demand or costs go up sharply, it might be efficient for the entrant to raise its prices. But it may not do so because going above a price 20% below the pre-entry incumbent price will free the incumbent from its own price floor. The entrant will thus bear some inefficiency in its pricing to get the benefits of imposing an inefficient price on the incumbent. From the incumbent’s perspective, the existence of the entrant is only one factor that might influence its pricing. To set a price floor at pre-entry levels

344. See *supra* Section III.A.

345. Edlin, *supra* note 6, at 982.

346. *Id.*

347. *Id.* at 945, 967.

ignores all the other reasons for lowering prices, like technological changes or drops in demand or costs. This will invariably produce inefficiency.

One need only recall all the distortions under Nixon's wage and price controls, which caused inefficiencies that took the rest of the decade to sort out. Or consider specifically the airline industry that provoked these proposals. There, costs routinely change sharply with shifts in fuel or labor costs and demand not only varies with economic cycles but predictably varies seasonally. Sometimes the shifts are even sharper. Imagine how disastrous it would have been to freeze airline prices right before the terrorist attacks of September 11, 2001, dramatically reduced demand for airline flights. Further, freezing into place inefficient prices on a route between the hub city and spoke city *A* not only causes inefficiency in that market, but spreads inefficiency to all the connecting flights from spokes *B-Z* to the hub that in part transport customers who travel on to spoke *A*.

Edlin attempts to address the problem of changing market conditions in two ways. First, he allows for an exception when after-entry costs fall "dramatically," which he defines as falling by at least 20%.³⁴⁸ But this does not alter the inefficiency of the price floor for any cost reduction below 20%, nor the inefficiency of the effective price ceiling on the entrant if costs increase. Nor does it alter the inefficiency of the price floor (and ceiling) if there have been changes in demand rather than cost. And it renders the Edlin restriction ineffectual whenever costs do go down by 20% or more.

Second, Edlin sets a twelve- to eighteen-month outside limit on his ban on reactive price cuts.³⁴⁹ But this does not eliminate the problem during that twelve- to eighteen-month period. Any changes in market conditions that do occur will make the short-term pricing freeze inefficient. Nixon's wage and price freeze, after all, only lasted three months.³⁵⁰ Moreover, setting the twelve- to eighteen-month outside limit only reinforces the long-run futility of the ban on reactive price cuts, and, if the moment of entry is defined to be when entry is first foreseeable to make the restrictions effective, this may mean the restriction expires before the entrant even sells its product.³⁵¹

The Williamson rule might seem to be a flat rule like the Edlin rule, only substituting pre-entry output for pre-entry price. But, seeing one of the problems of changing market conditions, Williamson recognizes that such a flat rule would be a disaster if demand increased. So, in the end, he

348. *Id.* at 970.

349. *Id.* at 945-46, 969.

350. See JACK E. MEYER, WAGE-PRICE STANDARDS AND ECONOMIC POLICY 67 (1982). The additional problems created by the more flexible wage and price controls applied in the months after the freeze was lifted give testament to the difficulty of making price adjustments based on changing economic conditions. *Id.* at 67-69.

351. See *supra* Section V.A.

proposes that the output ceiling be “demand adjusted.”³⁵² But this creates all sorts of new problems. Just how is one supposed to know how much of an adjustment in output to make when demand has increased? Williamson tries to get around this problem in various ways.³⁵³

First, Williamson suggests projecting future demand from past trends. But there is no reason to think this chartist approach works any better for projecting demand than for predicting future stock prices. Demand goes up and down depending on changes in consumer income, preferences, innovation, prices, and quality, as well as on the availability, price, and quality of substitutes. Courts cannot accurately project future demand from a past trend. Nor will that inquiry give an unambiguous answer since every trend will depend on the dates one picks and adjustments one makes.³⁵⁴ In any event, the extent to which increased demand will justify increased output depends on the intersection of that new demand curve with the incumbent’s cost curve. Thus, adjusting for demand cannot avoid the problems of inquiry into costs; rather, it multiplies them by requiring inquiry up and down the cost-output curve. Perhaps most worrisome, limiting future output based on past demand trends discourages incumbents from making investments in innovation and product improvements designed to accelerate any trend of increased demand.

Second, given inaccuracies in trend projection, Williamson changes his test to allow an output increase up to 10% above the projected demand. But this 10% buffer makes his restriction ineffective at protecting entrants (and thus unambiguously harmful) when demand has not increased by that amount. Further, it has the same flaw as the U.S. Departments’ approach. It does not eliminate the ambiguity; it just moves the ambiguity to a new point. Now the ambiguity will be about whether or not the incumbent is at a point 10% above an ambiguous demand-adjusted output. These ambiguities are worsened if, as Williamson did in response to criticism, one varies the percentage buffer from case to case based on the estimated degree of projection error.³⁵⁵

Third, Williamson says that when predatory pricing is alleged in one of many multiple geographic markets, then a simple comparison will tell us whether output in one of those markets has increased “disproportionately.”

352. Williamson, *supra* note 7, at 305-06, 333-34.

353. *Id.* at 305-06.

354. Williamson suggests relying on internal incumbent records. But since incumbents do not have a crystal ball either, they will often err in their projections. This is not so costly when firms can adjust to future realities, but if firms are bound by projections, even when erroneous, the costs are much higher. It is not clear why we want to visit such high antitrust penalties on those projections that do not pan out. Further, any rule based on incumbent documents simply invites the strategic drafting of documents in response to the rule. Although Williamson assumes firms will set actual output and prices in response to legal rules, he is oddly dismissive of the notion that they will take the less costly tack of changing the wording of their documents in response to legal rules. *Id.* at 305-06, 333-34.

355. Williamson, *supra* note 80, at 1192 n.40.

Obviously, that only applies when the incumbent is in multiple geographic markets and has monopoly power only in some. And even when it does apply, the ambiguities remain great. It will generally be unclear whether one can properly analogize between demand in different geographic markets. They have different consumers with different consumer preferences and incomes. Quality might differ. The markets might have different input costs, or different degrees of market power, both of which influence prices and thus affect realized demand for the good. Likewise, substitutes might differ in price and quality, and often their availability will differ because some geographic markets are further from substitutes than others. There is thus no general reason to think that demand will rise by the same amount in different geographic markets, or that courts can accurately quantify the differences. And all these ambiguities are exacerbated if the court is asked whether the difference is "disproportionate," which will mean varying things to varying judges and jurors.

True, these problems are somewhat reduced because, like Edlin, Williamson sets a twelve- to eighteen-month limit on his rule. But this does not eliminate the problem during that period, and it reinforces the long-term futility of the restriction. There are also other problems. Although Williamson adjusts for an increase in demand, he makes no adjustment for a decrease in costs, even though that too would indicate the efficiency of an expansion in output. Furthermore, where a product is variegated or changing, it may be difficult to determine what even the baseline pre-entry "output" was. The Williamson rule raises particularly difficult problems when a firm responds to an output ceiling by introducing a "new" product that is similar to the old product, but varies from it somewhat.

But the problem is not with these particular adjustments. The problem is that no effort to tweak these restrictions can eliminate the underlying problems. Those problems are rather an inherent consequence of trying to regulate incumbent pricing or output. There are two basic methods of implementing such regulation.³⁵⁶ One can, like Edlin, use a bright-line rule that, as stated, is over- and underinclusive and thus sacrifices facial correlation to the factors that affect its social desirability in light of changing market conditions. Or one can, like the European doctrine, the U.S. Departments, and Williamson, use a standard that correlates better to such social desirability criteria but cannot be applied as precisely, and thus will also be over- and underinclusive in actual application. Whichever method one chooses, the regulated price will often fail to reflect changing market conditions accurately and thus produce additional inefficiencies.

356. See Stephen McG. Bundy & Einer Elhauge, *Knowledge About Legal Sanctions*, 92 MICH. L. REV. 261, 267-79 (1993).

D. *Conclusion on Implementation Difficulties*

The fundamental problems posed by restrictions on above-cost pricing are exacerbated by other problems whose precise nature varies with the specific restriction but that cannot be avoided in one form or the other. The moment of entry that begins the period of restriction will either be defined in a way that makes the price restriction ineffectual, or in a way that makes the period of price restriction longer than the period of entrant pricing, exacerbating its adverse effects on pricing and innovation. Quality changes will either be left unregulated, which makes the price restriction even more ineffectual and encourages inefficiently high levels of quality, or will also be restrained, squelching desirable innovation. The price floor will either be fixed, freezing prices at levels that become inefficient as market conditions change, or uncertain, causing similar inefficiencies because of imprecise application and driving up prices because of risk aversion. All these are serious adverse effects that make it even less likely that the restrictions will have beneficial effects.

VI. THE BAUMOL BAN ON IMPERMANENT REACTIVE PRICE CUTS

Professor Baumol offers a somewhat different rule from those that would prohibit certain reactive above-cost price cuts. He would allow an incumbent monopolist to make reactive price cuts, but forbid those reduced prices from being raised after the entrant leaves the market unless costs or demand have changed.³⁵⁷ He would apply this price ceiling for a quasi-permanent period, and suggests five years as a possible choice.³⁵⁸ This rule would not make any reactive above-cost price cuts themselves illegal, but the Baumol rule would mean that making a reactive above-cost price cut subjects incumbents to a regime of price regulation that itself imposes costs on them. In particular, it forces incumbents to keep in place a price that might become less profitable if antitrust courts do not correctly adjust for changes in market conditions. Baumol's rule thus amounts to a restriction on reactive above-cost price cuts with a unique penalty. The penalty would not be standard antitrust damages. The penalty is instead whatever costs are associated with triggering the equivalent of quasi-permanent monopoly rate regulation.

Edlin argues that the Baumol rule should be rejected because it does not fit the standard *Grinnell* definition of prohibiting conduct that tends to create or maintain monopoly power by excluding rivals.³⁵⁹ Instead, Edlin

357. Baumol, *supra* note 7, at 4-6.

358. *Id.* at 8.

359. Edlin, *supra* note 6, at 978.

argues, the Baumol rule prohibits a price increase that, if anything, would encourage entry that might end the monopoly power. But the Baumol rule does not really prohibit price increases simpliciter. It prohibits impermanent reactive price cuts. Thus, if its effects were desirable, one could easily square the Baumol rule with the standard legal definition. One need only say that impermanent reactive price cuts are *not* deemed "competition on the merits," but rather are deemed strategic anticompetitive pricing to exclude rivals, whereas quasi-permanent reactive price cuts *are* deemed "competition on the merits" since they only drive out entrants through means that confer enduring benefits on consumers. As usual, whether or not we treat the conduct in question (an impermanent reactive above-cost price cut) as "competition on the merits" must turn not on conclusory legal labels but on a close analysis of whether banning that conduct on balance has desirable consequences.³⁶⁰ Once that analysis is completed, the legal label should follow.

A. *Post-Entry Effects*

Post-entry, there are two possible scenarios. One possibility is that the costs of triggering quasi-permanent rate regulation will be sufficiently high that the incumbent will be deterred from cutting prices as much as it otherwise would have. In this case, the Baumol rule effectively sets a post-entry incumbent price floor, and the effects are the same as considered in Part IV.

The other possibility is that the cost of triggering rate regulation will not deter the monopolist from cutting prices.³⁶¹ In this case, the long-term post-entry effects might be favorable, because, after the entrant has exited, the monopolist will have to keep prices at that lower level for some quasi-permanent period. Thus, one cannot say of the Baumol proposal, as one can of the others, that it is necessarily futile in the long run. This apparent advantage is, however, more than compensated for by the fact that such a long-term price ceiling creates even worse implementation difficulties and adverse incentive effects.

B. *Implementation and Incentive Problems*

To avoid problems in defining entrants, Baumol ultimately triggers his rule by exit rather than entry. His price ceiling applies "to any firm whose low prices are *suspected* of having driven its competitor from the field,

360. See *supra* Section I.B.

361. See Edlin, *supra* note 6, at 978-79.

whether or not that competitor was a recent entrant.”³⁶² But this exit test raises many new problems. First, the fact of exit can be unclear or invite strategic manipulation. What happens if a price cut does not drive a rival out of the market but reduces it to a crippled fringe size? If that does not count as an exit because the firm is still “in the field,” then an incumbent will have incentives to inefficiently decline to service some set of customers in order to leave entrants in business. And if small entrants do not count, courts have to define just what the size threshold is.

Second, the cause of exit will often be unclear and yet so plausibly connected to rival price as to make the Baumol rule ubiquitous. Firms exit markets all the time. Their exits have multiple causes that are difficult to sort out, an uncertainty only worsened by a test based on whether a causal link to the price cut is “suspected.” Indeed, failed firms always could plausibly connect their exit to their rivals’ prices. After all, presumably at some price they would have stayed in the market. Do we really want every firm exit to trigger rate regulation of any remaining firms in that market that have market power?³⁶³ That undermines normal market competition since, in most cases, firms have market power precisely because they are more efficient and thus able to charge lower prices than their rivals.

Third, even if we know we have a qualifying exit, we must define the precise moment of exit that determines when, and at what price, the cap is triggered. What happens if the incumbent increases prices just before the entrant exits? Baumol allows the incumbent to rescind a price cut if the entrant is still “alive and well,” but that raises difficult questions about just how well the entrant has to be.³⁶⁴ In practice, there will be varying prices during any period of incumbent-entrant competition. It will be unclear what time to use as the baseline, and choosing any particular time invites strategic manipulation.

Even if exit issues are resolved, defining the price and product on the exit date can be hard when both are variegated, and when associated terms influence the effective price. The incumbent also has incentives to introduce a related “new” product to evade the ceiling,³⁶⁵ requiring an unwelcome choice between allowing evasion and clamping down on new innovations. Further, after that date, demand shifts will require changing the price ceiling, with all the problems described above for the Williamson

362. Baumol, *supra* note 7, at 6 n.17 (emphasis added).

363. Baumol does not limit his proposal to monopoly cases but includes any oligopoly market in which firms have market power. *See id.* at 5 n.15.

364. *Id.* at 4 n.12.

365. Baumol would prohibit withdrawal of the old product, *id.* at 9 n.24, but that raises difficult enforcement problems. How are courts to decide how many sales of the old product must be made, whether sales and advertising have become insufficiently enthusiastic, or whether service and delivery have become too surly?

approach.³⁶⁶ Likewise, cost shifts will require changes with all the problems of rate regulation noted for the profit-maximizing price floors.³⁶⁷ But the problems are even worse. If the price ceiling is mistakenly set too low, it can make the incumbent lose money and even drive it out of business entirely. Further, if a mistakenly set price ceiling reduces incumbent output, by hypothesis there will be no entrant to take up the slack in output.

Finally, trying to maintain the price regulation for a quasi-permanent period of five years exacerbates the problems created by changing market conditions. True, one could try (and Baumol is open to) other specifications of the period of price restraint.³⁶⁸ Professors Joskow and Klevorick, for example, basically adopt the Baumol rule but change the period of quasi-permanence to two years.³⁶⁹ But the quicker the price ceiling expires, the more ineffectual the rule. Thus, the underlying problem remains that, no matter what specification one makes, one faces the problem of greater inefficiencies the longer the period is and greater ineffectualness the shorter the period is.

To try to get around these line-drawing problems, Baumol allows price increases as long as they are within an "order of magnitude" of the claimed increase in demand or costs.³⁷⁰ By now this gambit should be familiar, and it has the same problem as the efforts to avoid line-drawing by saying a price or output has to "clearly," "substantially," or "disproportionately" exceed some benchmark.³⁷¹ All these rules move the ambiguity to a new point but cannot eliminate it. And they do so at the cost of making the posited rule ineffectual. Baumol presumably does not mean the mathematical definition of an order of magnitude since that would allow any price increase as long as it was within a multiple of ten of the posited increase in demand or costs, and would really make the rule ineffectual. But whatever meaning is given to the term, some trade-off of harmful effects remains.

Where the Baumol price ceiling most exacerbates the types of concerns considered above is in its adverse effects on innovation. Like the price floors, a price ceiling will induce quality changes to evade the price restrictions that are inefficient and would not otherwise have been tried. But now the incumbent can be expected to try to evade the price ceiling by

366. See *supra* Section V.C.

367. See *supra* Section V.C. The Baumol rule does not set a ceiling at the profit-maximizing level but raises similar problems, because it effectively sets a ceiling at whatever price level creates the same profitability as the reactive price cut.

368. Leaving the period defined as "quasi-permanent" would be utterly vague and would worsen uncertainty problems.

369. Joskow & Klevorick, *supra* note 71, at 255 (applying it to reactive price cuts above average total cost).

370. Baumol, *supra* note 7, at 7.

371. See *supra* Section V.C.

making its product *worse* through cheaper production, so that it can still earn a monopoly profit, which not only creates an inefficient price-quality trade-off but degrades product quality. Even more problematic is the case where the price ceiling cannot effectively be evaded. Then, any investment in innovations to improve the product in a way to make it more valuable to consumers will be discouraged because the incumbent will not be able to raise prices to reflect that extra value and recoup the cost of that investment. Such a lowering of productive innovation is likely to be far more detrimental than any gain in allocative efficiency.³⁷²

C. *Ex Ante Effects*

The Baumol ban on impermanent reactive above-cost price cuts, if anything, will offer even less encouragement to entry than the various restrictions on reactive above-cost price cuts. If entrants foresee that the rule will not prevent incumbents from cutting prices to a level that will drive them out, the rule cannot encourage entry at all. If entrants instead foresee that the rule will effectively impose a price floor on the incumbent, then it will have the same effect as the proposals of Edlin, Williamson, or the Departments. The long-term futility of protecting less efficient firms means that their entry will hardly be encouraged, and the rule provides no encouragement (and possibly some discouragement) to the more efficient firms that would otherwise enter.

The Baumol rule is even less likely to create incentives for ex ante limit pricing than the other rules because it offers less encouragement to entry. Indeed, since the incumbent retains the option of driving out the entrant with a reactive price cut that amounts to an ex post limit price, it is hard to see why the incumbent would ever adopt that limit price ex ante.³⁷³ They would be better off charging a monopoly price and imposing a limit price only for a quasi-permanent period after entry, rather than charging a limit price every day. Not only would the reactive strategy mean that incumbents would get to charge a monopoly price rather than a limit price on more days, it also means that incumbents are less likely to impose a limit price that is unnecessarily low because incumbents imagine entrants might be more efficient than they turn out to be. Instead, incumbents can impose just the right post-entry limit price to drive out the entrant.

Finally, when entry occurs, and incumbents respond with price cuts that trigger a long-term price ceiling, the Baumol rule discourages innovation and investments by the incumbent in product improvement or, even worse,

372. See *supra* Subsection IV.C.1.

373. Edlin reaches the same conclusion that the Baumol rule will never induce ex ante limit pricing, but does so based on different reasoning. Edlin, *supra* note 6, at 979.

encourages product degradation. The problem is not just that this will occur post-entry, but that the prospect of such an *ex post* restriction on incumbent prices will reduce each firm's *ex ante* incentives to make the investments of time and money that created something so valuable it enjoyed monopoly power.³⁷⁴

VII. CONCLUSION

Both recent and longstanding analysis supporting a ban on above-cost predatory pricing requires a better response than current scholarship has so far given for why predatory pricing should be restricted to below-cost prices. This Article endeavors to provide that response by showing why efforts to restrict above-cost reactive price cuts are likely to be futile and harmful.

One reason is that reactive above-cost prices often do not protect incumbent market power at all, but rather are a normal competitive response to an entrant who has undermined a competitive schedule of discriminatory prices that maximized total output given common costs. This is particularly true in the airline industry, which provoked the recent proposals to restrict reactive above-cost price cuts. Airlines that operate hub-and-spoke systems incur common costs that cut across their different routes, and although on some routes their market share and prices may appear high, they are in sufficiently vigorous competition with other airlines that they earn no positive economic profits. Thus, it is likely that their price discrimination among different routes reflects not market power and an ability to reap supracompetitive profits, but rather their competitive adoption of the schedule of discriminatory prices that maximizes the total output of their hub-and-spoke systems. Reactive price cuts in response to entry that undermines that output-maximizing schedule of prices is a natural competitive response, and efforts to prevent such price cuts would likely have the adverse consequence of increasing prices along the rest of the hub-and-spoke system, reducing the total output of the hub-and-spoke system, and harming aggregate consumer welfare.

Even when incumbents do have market power, restrictions on their ability to adopt reactive above-cost price cuts are unlikely to achieve the objective of encouraging and protecting entry because less efficient entrants cannot survive in the long run, and entrants who are (or will predictably become) more efficient need no encouragement or protection. Further, such restrictions will have harmful effects by raising prices and lowering productive efficiency during any period of price restriction, inflicting wasteful transition and entry costs, as well as distorting innovation and price flexibility in response to changing market conditions. And the

374. See *supra* Section IV.D.

restrictions will discourage the creation of more efficient incumbents and entrants, which is ultimately far more important.

This analysis reaffirms the wisdom of the position that antitrust law should not recognize any claim of above-cost predatory pricing. It also helps specify just what should count as costs. Costs should be defined in whatever way assures that an incumbent pricing at cost could not deter or drive out an equally efficient entrant. This test should be met by a cost measure that includes all costs that are varied by the allegedly predatory increase in output, since short-term threats or pricing strategies that exceed short-term costs should not be able to deter long-term investments or entry. Alternatively, if short-term pricing could deter such long-term decisions, this test would be met by a cost measure that reflected the magnitude of predator costs for the sorts of costs that are variable to the rival during the period of entry or investment decisions influenced by the short-term existence or threat of such pricing.
