Journal of Regulatory Economics; 21:1 35–56, 2002

The Deregulatory Tar Baby: The Precarious Balance Between Regulation and Deregulation, 1970–2000 and Henceforward*

ALFRED E. KAHN
Cornell University
Robert Julius Thorne Professor of Political Economy, Emeritus, Special Consultant, National Economic Research Associates, Inc. (NERA), 308 North Cayuga Street, Ithaca, NY 14850
E-mail: alfred.kahn@nera.com

Abstract
The recent history of the airline, telecommunications and electric utility industries years amply demonstrates the benefits of deregulation. It also demonstrates both the necessity of continued government involvement and its pitfalls. Prominent among the former are the efficient provision and pricing of infrastructure services, enforcement of antitrust-like policies and regulation of bottleneck facilities. The latter come down to an inherent tendency of regulators to engage in continued micromanagement—to continue to assume responsibility not merely for opening the door to competition but for ensuring that competitors go through it, and to prescribe the results they think an efficient competitive process would produce.

1. Introduction

When I became chairman of the New York Public Service Commission, in the summer of 1974, I had not the slightest notion that I was assuming the leadership of an enterprise that either was or should have been in process of going out of business. True, I soon discovered that my distinguished predecessor, Joe Swidler, had agreed with Rochester Telephone to undertake an experiment of opening the provision of terminal equipment—phone instruments, answering machines, PBXs—to competition, subject only to certification of compatibility of competitive equipment with its own network: prevailing policies still permitted the franchised local monopolies to require (and charge prices far above cost for) ''protective connection devices'' between ''alien attachments'' and the interface with the telephone network.[1] The purpose of the experiment was to see whether, as some exuberant Bell representatives claimed, removal of that protection might end up electrocuting

* I thank Dennis Weisman and Tim Tardiff for their extremely careful reading of this manuscript and suggestions for its improvement.

1 See the discussion and criticism of the Hushaphone and Carterphone decisions in Faulhaber (1987, 27–30) and in Kahn (1971, vol. 2, 140–145).

telephone linemen. The notion, however, that other parts of the business could be opened to competition or deregulated still seemed remote. This was true even of long distance service, despite the FCC's decision in 1959 to permit use of the spectrum above 890 Mc for private microwave systems, its interminable proceedings thereafter concerned with defining the acceptable response by AT's to that competition, and its 4–3 decision in 1969 to approve MCI's provision of private carriage alone.[2] The only major active deregulatory issues at the national level related to the field price of natural gas, stock exchange brokerage commissions and whether cable systems should be relieved of the prohibition of their importing over-the-air distant signals, imposed in the interest of protecting local broadcasters (Kahn 1971, vol. 2, 193–209 and Noll et al. 1973, chapter 6).

My preponderant interest, instead, was in applying to electric, gas, telephone and water system charges the elementary economic principles I had previously enunciated in my first volume.[3] Meanwhile, the Kennedy hearings on the airline industry, directed by a Harvard law professor, Steven Breyer, were only just getting underway. Although I testified at them, I was of course totally unaware of how, within fewer than three years, they would change the course of my life.[4]

2. The Continuing Tension Over the Proper Role of Government in the Deregulated Airline Industry

It seems fitting to use the airline industry as my first illustration of ''the precarious balance between regulation and deregulation.'' I have elsewhere described how what we intended to be a gradual process of deregulating the airlines soon took on a life of its own, like the proverbial snowball rolling down a hill—the mirror image of the tendency of regulation, once undertaken, to become increasingly pervasive and thoroughgoing.[5] Unsurprisingly,

---

2 See Kahn (1971, vol. 2, 129–136). Also Faulhaber, chapters 2 and 3. MCI's defiance of the Commission's prohibition of its offering public switched service was not sustained in the courts until 1977 (MCI v. FCC, CCA, DC Circuit, July 28, 1977).

3 See the ''Introduction: A Postscript, Seventeen Years After,'' to the 1988 reprinting, especially its survey and interim assessment of ''The Deregulation Revolution,'' and ''The Anomalies and Distortions of Partial Deregulation.''

4 This is an apt occasion to acknowledge my debt to Walter Adams for calling my—and the world's—attention 20 years earlier to the similar cartelization of the trucking industry, largely in his work for the Senate Small Business Committee. See the discussion and references *ibid*., vol. 2, esp. 14–18 and 186–193.

5 See the reference to Kahn 1971 in note 3, above, a discussion that documents the ''inherent tendency to what might be termed service inflation, in which an equilibrium of cost and price is achieved not by reducing price to marginal cost but raising marginal cost to price'' (207).

> The typical answer of foreign governments to asserted excessively low load factors, namely the imposition of direct limitations on the amount of capacity offered, market by market, provides yet another illustration of the inexorable tendency for regulation of a competitive industry to spread. Control price, and the result will be an artificial stimulus to entry. Control entry as well, and the result will be an artificial stimulus to compete by offering larger commissions to travel agents, advertising, scheduling, free meals, and bigger seats. The response of the complete regulator, then, is to limit

therefore, when and as deregulated airline markets began to demonstrate the consequences of severe imperfections—

1. $13 billion or so of losses industry-wide during the 1990–1993 period—greater, the industry proclaimed, than the totality of its profits since the Wright Brothers' historic achievement in Kitty Hawk, North Carolina—demonstrating all of the symptoms of destructive competition, consequent on the splurge of capital expenditures in the late 1980s and the subsequent deceleration in the growth of demand;
2. the general deterioration in the quality of the air travel experience and especially extreme congestion at particular times and airports in very recent years; and, finally,
3. the extreme profusion and differentiation of fares, reflecting in important measure differences in marginal costs but also genuine price discrimination, reflecting some degree of monopoly power, particularly at increasingly prevalent hubs dominated by single carriers—it predictably and understandably generated increasingly intense pressures for governmental intervention. I confine my attention to the last two; the first has for obvious reasons come to seem much less urgent than it was a decade ago.

In principle, those proposals have fallen into two very different categories—some (coming particularly from parties who were never enthusiastic about deregulation) clearly constitute proposed reversals of the historic decision in 1978—putatively partial but, all experience tells us, certain to be driven in the direction of totality—others, at least in intent, involving no more than an application of antitrust principles to preserve the competition that is the essential substitute for regulation.

But of course, as every student of antitrust knows, that important distinction, simple and unexceptionable in principle, nevertheless in large measure merely shifts the nominal terms rather than the substance of the debate: to what extent do proposed interventions, under the rubric of preserving competition—the very rationalization of comprehensive regulation in the first place—actually involve repressing it in the interest of protecting competitors?.

I propose to offer two examples, knowing that for this audience little elaboration will be necessary.

2.1. Quality of Service Generally and Congestion in Particular

About the deterioration in the quality of airline service generally—the longer lines, more crowded planes—thanks largely to the 20% point increase in average load factors and narrower seating—I can be very brief. Under tight price regulation, travelers were confronted with only one choice—generally good (indeed, inefficiently good[6]) service but

---

advertising, control scheduling and travel agents' commissions, specify the size of the sandwiches and seats and the charge for inflight movies. Each time the dyke springs a leak, plug it with one of your fingers; just as a dynamic industry will perpetually find ways of opening new holes in the dyke, so an ingenious regulator will never run out of regulatory fingers (1978b).

6 See footnote 5, above.

at cartel prices; the purpose of deregulation was, by unleashing price competition, to offer economy-minded travelers an alternative low price and necessarily correspondingly lower quality service; and in that it has dazzlingly succeeded.

In contrast, the intolerable congestion at major airports, either at particular times or chronically, is probably the greatest single black mark on the industry's performance. The temptation is to characterize it, correspondingly, as the single greatest failure of deregulation.

Obviously, it is not the latter: on the contrary, it is the strongest possible vindication of the expectations of the deregulators—expectations of an accentuation of price competition and an elastic response of the traveling public to the sharp consequent reductions in average fares. The failure—the only failure—was that of governmental agencies to respond to the predictably increased demand for infrastructure—specifically, air traffic control and airports—and to price it correctly—a blame that I can definitively deflect from myself (and from deregulation itself), since I explicitly emphasized to the FAA the importance of their fulfilling that function at the very time we were undertaking to deregulate (Kahn 1978a).

I do not have to make the case here that the efficient remedies are and always have been reform of our organizations for providing those infrastructure services—removing them from the regular agencies of government and the federal budget, whether by corporatization or privatization—and the introduction of marginal cost pricing by a combination of congestion charges and auctioning of slots.[7] The only amplifications worth making in the context of this talk are to rebut the naïve assumption—of which surely none of you will be guilty—that congestion charges will produce an additional burden on airline customers; to document the danger that this abominable performance may lead instead to ''solution'' by reregulation, rather than by efficient pricing; and to consider the ways in which antitrust protections may be required.

Consumer organizations rarely put efficient pricing high on their list of proposed remedies, evidently partly for fear that the result will be higher prices for scarce services. But average fares are, unsurprisingly, already higher on flights to and from slot-constrained airports and on peak than off-peak, reflecting the comparative availability of discount tickets. Efficient congestion-based charges—or large proceeds from slot auctions—at times of excessive congestion will extract those rents from the carriers, rather than raise fares. Small wonder the favored carriers oppose them, preferring instead to assume responsibility for tailoring operations to capacity cooperatively, under grant of antitrust immunity![8] The other side of the coin would have to be arrangements assuring that those rents would, in turn, be used for some combination of increasing capacity and subsidizing—in the form of negative charges, if necessary for revenue neutrality—usage at off-peak times or uncongested auxiliary airports. Efficient charges of this kind would further realize the success of deregulation itself (and, incidentally, also of the bumping rules enacted by the CAB) in offering travelers a low-fare/low-quality (or lower-convenience) option denied them under regulation.

---

7 See, for example, Levine (1969).

8 For a thorough criticism of this proposed ''solution,'' see Robyn (2001).

As to the second, it is difficult to do justice, summarily, to the flood of proposals for legislative enactment of ''travelers' bills of rights'' engendered by the clearly unsatisfactory air service to which they have been exposed in recent years, particularly when, at peak travel times, adverse weather conditions bear principal responsibility for the most dramatic deteriorations. As the one who, inspired by the late Julian Simon, persuaded the CAB to adopt its present bumping rules under our authority to prohibit unfair or deceptive methods of competition, I would be among the last to oppose legislation banning ''deceptive scheduling of flights,'' requiring the publication of on-time performance or requiring (or, as I would prefer, requiring some administrative agency to require) compensation of aggrieved passengers in the event of specific failures. Yet the tendency of consumerists to prescribe such rules; to assign the principal blame to the scheduling practices of the carriers—as though they have any choice individually but to schedule their flights, to the maximum extent possible, when and where customers want to travel—rather than to inefficient pricing of infrastructure services; to prescribe preferential allocations of scarce facilities to favored carriers—low-fare or regional; indeed the hostility some of their proposals betray to efficient pricing ''as a way for the FAA to raise additional revenues rather than reducing congestion delays,'' to quote one of them—all reflect a bias in favor of regulatory rather than efficient competitive solutions.

There remains to be considered the possibility that the use of congestion pricing or slot auctioning may be distorted by market power. This is clearest in the case of auctioning: the value of a slot is likely to be systematically higher for the bidding carrier for whom it preserves a dominant position and the accompanying monopoly power than for the carrier to whom it offers only an opportunity to compete with an entrenched rival. The response, in principle, would be that it is precisely the purpose of the antitrust laws, and of section 7 of the Clayton Act in particular, to prevent such agglomerations of assets. Enforcement of that protection through the regular antitrust channels—involving ultimately adjudication by the courts—is both highly uncertain (for good economic reasons, no doubt) and cumbersome. For that reason, the antitrust agencies themselves use essentially administrative means not only, necessarily, to decide what mergers they will attack but, with increasing frequency, to attach structural conditions—typically stipulated divestitures—to an agreement not to prosecute. So the most efficient way of counteracting any such distortion of the bidding process would be to set some limits—say, 20 or 25%—to the share of the total slots for which any carrier could successfully bid.

Without some such protection, congestion pricing would seem to be subject to the same distortion as slot auctioning: as the price at times and places of excess demand rises to whatever level necessary to reduce congestion to optimal levels, presumably those carriers for whom use of the facilities protects a monopoly or dominant position would be the ones who would not drop out. The only answer I can think of would be to impose a similar limitation on the permissible share of any single carrier. To the foreseeable reaction that such a preventive would be inherently regulatory, the answer can only be that it is no more so than section 7 of the Clayton Act and the spirit of the antitrust laws generally, the importance of which is of course enormously increased by the abandonment of direct regulation itself.

These suggestions raise the familiar dilemmas of all such limitist solutions to antitrust issues—in the present instance, the fact that hubs have characteristics of natural—as well

as unnatural—monopoly[9] and that, as Dan Kasper has urgently called to my attention, such limits could well weaken the important competition on long routes among network carriers over their respective hubs.[10] The dominant carrier may well be able to outbid its rivals for the scarce slots or better pay the congestion-based fees because of the superior economies of scope that usage of airports at those times permitted it to exploit.

2.2. Extreme Price Differentiation

> It costs $352.50 for a round-trip coach ticket between Boston and Washington, a distance of 406 air miles. But if you want to go to London, 3,267 miles away, it costs as little as $298—provided you meet the restrictions.[11]

In my response to this complaint, I compared the $352.50 full Boston/Washington fare with the weighted average round trip fare for that route of $226.22, and the $298 Boston/London fare with the range of $401, $506 and $1,587 Boston/London, respectively, for the excursion fares, low and high seasons, and the regular tourist fare.[12]

These differentials are, of course, in important measure not discriminatory at all, but reflect differences in marginal costs; they are in large measure also definitely discriminatory but welfare-enhancing, in an industry characterized by extreme economies of scale and, even more, scope—benefiting the ''victims'' as well as the beneficiaries; but they create a need for the best possible guarantees that the demand-inelastic customers are in fact not being exploited—that is, are not being forced to pay more than the costs of serving them on a stand-alone basis (Baumol et al. 1988, 508).

There are only two generic ways of attempting—and of course the word ''attempting'' requires multiple emphasis—to provide such protection, regulatory and competitive. The first could involve ceilings of estimated stand-alone costs or fully allocated costs or setting some arbitrary maximum to the spread between average and full fares—all of which have been advocated at one time or another. The other would require doing everything possible to strengthen the discipline of competition: perfect contestability is the perfect preventive of rates to any group of customers exceeding stand-alone costs (Baumol et al. 1988, 352–354; 508–509).

My rejection of the regulatory solution is not merely aesthetic or an attempt to exorcise the hobgoblin of inconsistency; it is informed rather by experience and a recognition of the logical and practical impossibility of partial re-regulation. As for the former, the contestant parties in proceedings attempting to set ceiling rail rates for captive shippers under the

---

9 Kahn (1993, 381–387). See also Levine (1987).

10 For example, if such restriction applied at ORD [O'Hare], both AA and UA would be severely handicapped in competing against carriers such as NW and DL whose competing hubs are relatively unconstrained. In addition, it is clear that hubbing carriers at efficient hubs typically use well over 25% of the operations, even at unconstrained hub airports. Thus, the proposed limits would effectively eliminate hubbing if applied to hub airports. Another reason to push hard for airport capacity expansion. (Personal communication.)

11 Kuttner (1988). Undoubtedly more extreme examples could be cited today.

12 Kahn (1988).

provisions of the 1980 Staggers Act have accepted the proposition that stand-alone cost ceilings are theoretically correct but almost hopelessly difficult to administer, with the result that the captive shippers clause itself has provided little protection and is only rarely invoked these days.[13] This experience has suggested to me and, I believe, to others, that something like 180% of actual long-run incremental costs—measured in an approximate way by the railroads' so-called Form A costs—would make much more sense, pragmatically. An additional reason for rejecting any genuine attempt to measure those stand-alone costs has been the bitter experience with the FCC's prescribed standard for pricing unbundled network elements provided by incumbent to competitive local exchange companies—the total long-run incremental costs of a hypothetical competitor writing on an almost completely blank slate and employing the most efficient available technology—rather than the actual LRICs of the incumbent companies—to which I turn in part B, below.

My second reason for rejecting the regulatory protection is the tendency to which I have already referred and that I have already documented in other contexts, for ''regulation, once undertaken, to become increasingly pervasive and thoroughgoing.'' (See note 5, above.) Despite the unfortunate tendency of the discomforts and inconveniences of the low-fare/low-quality service the emergence of which, as I have argued, is one of the proudest achievements of deregulation, to spill over and sully the service of the travelers who pay full fares, the airlines have over the last several years been attempting to differentiate the quality of the two—with automatic upgrades to first class, separate lines at the check-in counter and extra frequent flyer awards for the full-fare passengers. Predictably, any attempt by regulators to put ceilings on those higher fares will have to be accompanied by corresponding, increasingly detailed floors under quality—the percentage of times in which the payment of full tourist fares is indeed accompanied by upgrades, the respective lengths of lines at the check-in counters, the quality of their food service: there is no halfway house.

This leads, as the only logical complement of deregulation, to greatly increased emphasis by both the antitrust agencies and the Department of Transportation—which has the same kind of authority to prosecute unfair or exclusionary methods of competition as the Federal Trade Commission has for industry generally—to protect and improve the conditions of competitive entry.

This means to me, in turn, that the great majority of economists who are of the opinion that airline deregulation has been a success and is worth preserving rethink the highly skeptical view of many if not most of them about the reality of predation.[14] All it takes, in

---

13 The major exception, I am informed, is for shipments of coal moving on essentially stand-alone railroads. See my citation of this experience in my recognition of the theoretical correctness of the stand-alone cost ceiling and my rejection of the FCC's similar prescribed blank-slate pricing of unbundled network elements in my *Whom the Gods Would Destroy* (2001, 61–62), and in part 2, below.

14 See my fuller and latest discussion of this issue (*ibid*., 36–39) and especially lengthy accompanying footnotes (66–70). The day before I first dictated this discussion, there appeared the following in the *New York* Times (2001):

> Early last month, Canada's Competition Bureau, a government antitrust watchdog, asked a federal court to order Air Canada to stop pricing fares below cost in Eastern Canada. In February, for

my opinion, is that DOT and antitrust agencies recognize the soundness of marginal opportunity (rather than only marginal production) costs as the theoretically correct floor of non-predatory prices[15] and solve the administrative (and essentially regulatory) pitfalls of a government agency attempting to apply such a standard[16] by leaving that responsibility to the responding carriers themselves, by requiring them, if their response is indeed followed by withdrawal of the challenger, to maintain their responsive levels of capacity and fares for some period of time—say two years. That would tend to ensure that an incumbent carrier would not lightly undertake a response that was in its own opinion profit-sacrificing (and therefore putatively predatory) in the expectation of being able to withdraw it if it succeeds in driving out the challenger; and, at the same time, give travelers the continuing benefit of the newly introduced competition for some substantial period of time, rather than permit its quick withdrawal.[17]

---

example, Air Canada, which had been charging about $400 for one-way fares between Halifax and Montreal or Ottawa, suddenly dropped fares to around $65, matching those of the six-month-old CanJet.

15 See references to the unanimous agreement with this proposition in the Report of the Transportation Research Board/National Research Council (1999), in my *Whom the Gods Would Destroy*, note 79 (69–70).

16 All the members of the afore-mentioned TRB/NRC committee explicitly recognized these difficulties, despite their roughly 50/50 division on the merits of the Department of Transportation's initiative to proceed against ''unfairly exclusionary'' pricing.

The recent decision of the U.S. District Court granting American Airlines summary judgment in the predation suit brought against it by the U.S. Department of Justice, finding that the plaintiff had failed to demonstrate that the airline's entry-responsive fares fell below average variable *production costs*, rejected the Department's proposed alternative standard—namely whether the carrier had by its particular response sacrificed profits that it might have earned by an alternative response—on the understandable ground that any attempt to determine whether ''the response at issue involved the carrier's forgo[ing] more revenue than it otherwise would have made . . . would be a litigation nightmare . . . and would invite excessive speculation about reasonable alternative strategies and the like . . .'' U.S. District Court for the District of Kansas, *Memorandum and Order*, 2000, *U.S.A. v. AMR Corporation et al.*, Internet version (2001, 49–51).

17 The District Court in the American Airlines case also committed the logical error—emphasized by its own use of italics—of defining successful predation as the restoration of pre-competitive entry levels only if and to the extent that those levels were themselves super-competitive:

> In sum, the government's evidence establishes only that, after decreasing during a period of low fare carrier competition, fares of four of the core routes. . .returned to approximately the same levels as before. *But there is no evidence that the prior fares were in fact supra-competitive*. (*ibid*., 57).

See also the conclusion on p. 62, for the same standard. If, however, one conceives of the function of competitive entry as one of subjecting those previous fares to a test of competition, clearly a highly discriminatory, pinpointed response that succeeds, by driving out the challengers, in restoring those previous levels will have frustrated that competitive test. The District Court's standard would force on the government the highly regulatory responsibility of determining whether those previously prevailing levels were indeed excessive. On this and other related tests, see Kahn (2001, note 79).

3. The Apotheosis of the Tar Baby: Telecommunications ''Deregulation''

Despite the difficult antitrust issues that they continue to raise, deregulation of the transportation industries—as well as crude oil, natural gas, stock exchange brokerage fees, cable television and wireless telephony—has been comparatively simple. It was comparatively easy for us to achieve the measure of success that I proclaimed some time after assuming chairmanship of the CAB-namely, ''the extent to which there would be no such agency when I was through'' or ''getting those eggs so scrambled that no one would ever be able to unscramble them.'' The most complete exemplification of the tar baby effect has been and continues to be telecommunications—25 years after the courts sustained MCI's defiance of the FCC's prohibition of it offering universal switched long-distance service and five years after Congress proclaimed the goal of deregulation. The reason, of course—both here and in electric power—has been the persistence of elements of natural monopoly that provided the justification—and/or rationalization (Gray 1942, 280–303)—for their treatment as public utilities in the first place.

Since I have—in some instances in collaboration with Timothy J. Tardiff and Dennis L. Weisman,[18] including substantial testimony on behalf of the local Bell telephone companies—sharply criticized at length the FCC's execution of its authority under the 1996 Act, it would come close to malpractice for me to use this occasion to repeat those arguments at any length. I will content myself therefore only with the following summary observations, first about its pricing prescription and second about its identification of network elements meeting the statutory standards for mandatory sharing.

As for the first, I find simply incomprehensible the Commission's adoption of the ridiculous blank slate version of TELRIC for the mandatory pricing of network elements—all the more so in view of the fact that I (sometimes in collaboration with Timothy J. Tardiff) and Jerry Hausman, separately, had emphasized directly to them the conceptual fallacy that it embodied. The literature had—to my satisfaction!—established the sufficiency of a requirement that they make network elements available at the actual long run incremental costs of the incumbents, with or without a markup approved by their state regulators sufficient to recover their historically determined revenue entitlement, subject to what I have termed the requirement of competitive parity and Baumol and Willig efficient component pricing (Baumol and Sidak 1994, 171–202, Kahn and Taylor 1994, 225–40).

Taken together with the Commission's totally expansive criteria for identification of the network elements subject to mandatory unbundling—reversed by a seven-to-one vote of the United States' Supreme Court—it is difficult to conceive of a standard more inherently contradictory of the Act's manifest and understandable desire to encourage competitive challenges by entrants constructing their own facilities than their ability to obtain such facilities instead from the incumbent at prices at the lowest level that they could conceivably achieve themselves.[19]

---

18 Kahn (2001, chapter 2); Kahn (1998); Kahn et al. (1999, 319–365).

19 Unsurprisingly, therefore, the only major facilities-based competition has taken place where prices were held by regulators outrageously above the incremental costs of incumbents and challengers alike in order

With Henry Higgins, I can only wonder or marvel ''what can have possessed'' the Commission's Eliza to have presumed in this way to stipulate what the *results* of the hypothetical competitive process would be: ''I cannot understand the wretch at all!'' The only answer I have been able to divine I have put in the title of my little book, *Whom the Gods Would Destroy* (2001).

My only other observation I claim entitlement to make as one of the few surviving students of Schumpeter. Whatever the case for the FCC's identification of ILEC network elements subject to mandatory sharing and its prescribed method of determining the price, it clearly becomes progressively attenuated the more such advantages as the ILECs enjoy from providing them are not simply the heritage of their previous franchised monopolies but flow from the undertaking of new, costly and risky investments, in an environment of technological and commercial uncertainty. Such a policy confronts investors with the asymmetrical prospect of being required to share the results of investments that turn out successfully, at regulated (let alone absurdly regulated) rates, while perforce absorbing the entire costs of the unsuccessful ones. The case turns unequivocally negative when, as the FCC has itself found with respect to the offer of broadband services generally and high speed Internet access specifically, the ''preconditions for monopoly appear absent''; the market is likely ''to accommodate different technologies such as DSL, cable modems, utility fiber to the home, satellite and terrestrial radio''[20]; and the market share of the phone companies is roughly half that of the cable companies, who are subject to no such sharing obligations.

In these circumstances, the FCC's insisting additionally on attaching to its authorization of the ILECs to do anything it is a position to deny—to merge, to offer interLATA services in region, to offer advanced, broadband services—no doubt to go to the bathroom if it could—its favorite condition, ''only if you accept the handicap of offering broadband services through a fully separated subsidiary''—seems, quite simply, anticompetitive.

Although as little as a year or so ago, I could not

> bring myself to deny the regulatory agencies a central role in the transition of public utilities to competition, in view of the special circumstances of those industries: the necessity for settling out and tracking the collection of strandable costs; the more pervasive possibilities in those industries of tying competitive to monopoly services, directly or subtly, and of cross-subsidization, strictly defined; the consequent need for

---

to cross-subsidize other services. The major historical instances of regulatory-mandated egregious overcharging were in (1) the retail charges for long-distance calling and then to independent long-distance companies for access to the local networks—leading to the emergence of independent long-distance access providers in every major metropolitan area in the United States—and (2) the retail charges for high-volume services to business users in those areas.

20 Report, *Inquiry Concerning the Deployment of Advanced Telecommunications Capability*, 14 FCC Rcd 2398, 2423–24, par. 48 (1999); *see also* Memorandum Opinion and Order, *Applications for Consent to the Transfer of Control of Licenses and Section 214 Authorizations from MedioOne Group, Inc. to AT&T Corp.*, 15 FCC Rcd 9816, 9866, par. 116 (2000); *see also* Seventh Annual Report, *In the Matter of Annual Assessment of the Status of Competition in the Market for the Delivery of Video Programming*, CS Docket No. 00–132, FCC 01–1, par. 52 (rel. Jan. 8, 2001).

> accounting separations and the monitoring of transactions between still-regulated utilities and unregulated affiliates; and, finally, the pervasiveness of essential facilities controlled by incumbents—not to mention competitive advantages deriving solely from their historical franchised monopolies, requiring an administrative agency to define them and to prescribe the terms and conditions of sharing (Kahn 2001),

I'm tempted to suggest a moral from the FCC's behavior: that the task of presiding over and establishing the necessary conditions for deregulating an industry should not be entrusted to the agency that regulated them.

4. The California Electric Deregulation Fiasco[21]

I had been promising myself during the last few months to determine a reasonable apportionment of responsibility for what has gone wrong in California better than the explanations one can read in any newspaper any day—to decide what conclusions can reasonably be drawn from it about the wisdom of that venture, both as a general proposition and in its major details, and to use this paper as the occasion for offering my insights to an eagerly expectant world. I have long since abandoned that ambition: there are too many others who understand far better than I the peculiar constraints imposed by the technology of electric power generation and transmission, and have articulated a far clearer vision of the *regulatory* policies and organizational arrangements essential to satisfy those constraints if competition is to be feasible.[22] Sheer cowardice, along with a recognition of my limitations, forces me to confine myself here to a series of questions and suggestions.

4.1. Should We Have Deregulated?

Although the question may be pointless at this stage—too many eggs have become scrambled; a generation sector that it still appears could be effectively competitive (even

---

21 My first and only exposure to the California experience resulted from my appointment by the California Power Exchange to the chairmanship of a ''Blue Ribbon Panel''—whose other members were Peter C. Cramton, Robert H. Porter and Richard D. Tabors—to assess proposals that the Exchange switch from uniform pricing to pay-as-bid. Our report, on January 23, 2001, has since been published in slightly altered form in the *Electricity Journal* (Kahn et al. 2001).

22 See Hogan 2001, in this issue, citing also the prescient discussion of Joskow and Schmalensee 1983; also Joskow 2000.

> The only involvements I have had with the radical restructuring of the electric utility industry have been testimony I have given on (1) the entitlement of the utility companies to recover the costs that, it appeared—indeed, it was fondly expected and hoped by some parties—would be left stranded by the introduction of competition and (2) the formulation of codes of fair conduct governing the anticipated competition at retail between the still-regulated local utility distribution companies or their affiliates and independent marketers. Even if the first of these appeared to involve many billions of dollars and the second was the subject of intense controversy, neither seemed to me of great importance compared with whether or how an electric industry, part deregulated and part, necessarily, still regulated, could work.

though it may not be in the present situation of extreme shortage) has now been separated out; any suggested reforms aimed at restoring regulation may well do more harm than good—it is still one that all of us must attempt to answer for ourselves. An answer would have to take into account, above all, the extraordinary and in some respects literally unique characteristics of this industry—the high degree of interdependence between investments and operations at the transmission and generating levels, hitherto presumably taken fully into account internally by the several vertically integrated franchised monopolists that dominated the industry[23] and by pools or bilateral arrangements that they found in their several interests to enter; and the especial need for integration created by the essential non-storability of power and the corresponding need for an instantaneous matching of generation and consumption, in which failure at any single point can threaten the stability of the entire interconnected grid. California has made especially clear how extraordinarily the effects of failures to satisfy these constraints are compounded by the highly inelastic demand on the part of most consumers, particularly because of the absence of any means of confronting them with the wide fluctuations in marginal supply costs on a real-time basis,[24] and the long-run inelasticity of supply—using long-run in both its popular, temporal, and technical economic sense.[25]

As a result, the industry is extraordinarily susceptible to severe consequences of forecasting errors. For at least the last 10 years, I have been pointing out that the deregulation movement in electric power has been essentially opportunistic: practically no one was calling for it during the 1950s and 1960s, when retail prices declined some 30% in real terms, or in the 1970s and early 1980s, when it appeared—and to this day in states where it still appears—unregulated prices would be higher than regulated; but deregulation became almost irresistible politically by the middle 1990s in those parts of the country where, it seemed certain, competitive prices would be far below regulated ones.[26]

While I, along with many others, recognized that the apparent opportunity for huge short-term benefits from deregulation was created by enormous forecasting errors in the 1970s—in assuming an indefinite continuation of demand growth at accustomed 6 or 7% annual rates; in estimating the costs of nuclear plants and in foreseeing what would in the 1980s and 1990s turn out to be the lowest-cost technology; in regulators' estimates of avoidable costs on the basis of which they were instructed by PURPA to set the prices the utility distribution companies were required to pay for independently generated power, as a result of all of which the industry, particularly on the East and West Coasts, found itself with what appeared to be a heavy load of high-cost excess capacity and a consequent wide spread between book costs and the estimated costs of new combined cycle plants powered

---

23 See, e.g., Kaserman and Mayo (1991, 483–502).

24 There is a story now to be told about the apparently unexpectedly large response of demand and its role in fending off the worst expected shortages in California in the summer of 2001.

25 See, among others, Rosen et al. (2000); and the brief, challenging ''Heresy? The Case Against Deregulation of Electricity Generation,'' by the pseudonymous ''Price C. Watts'' (2001, 19–24). See, in contrast, the confident exposition by Hogan (2001) of the policies of a still-regulated, naturally monopolistic transmission system sufficient to ensure efficient, socially beneficial competition in generation and distribution. Also Rowe et al. (2001).

26 See also Joskow (2000, 133–138, 142–143).

by unexpectedly cheap natural gas—the irony is that we failed to recognize that the generally accepted expectation of huge benefits for consumers from deregulation might similarly prove to have been based on the opposite forecasting errors. This time, it has been the growth in demand that has been underestimated; the persistence of excess capacity overestimated—especially when generators were vowing not to make their previous mistakes again!; the cost of natural gas and NO_X emissions underestimated; and the long-run inelasticity of supply either ignored—because of its seeming irrelevance—or its importance underestimated; and the importance of real-time marginal cost pricing and the short-run demand elasticity of which it is a prerequisite inadequately appreciated, despite our belief that we had learned that lesson in the inflationary 1970s.

In these respects, I have the clear impression, California was not unique: its experience was simply more extreme than elsewhere. In these circumstances, however, it is not surprising that observers in California seem to have been first to identify apparent exercises of market power because of the newly-discovered susceptibility of its spot markets to unilateral withholding by multi-plant producers of some comparatively small proportion of their capacity and the consequent apparent insufficiency of competitive disciplines at times of peak demand.[27]

4.2. Specific Errors or Uncertainties, in Hindsight

4.2.1. The Retail Price Freeze

Postmortems on California, many of them justifying Harold Demsetz's acute observation many years ago that ''believing is seeing'' (Demsetz 1974)—brought to my

---

27 Borenstein and Bushnell (1999); Borenstein et al. (2000); Joskow and Kahn (2001). Our Committee (see note 21, above, recognized the possibility that the Power Exchange's uniform pricing system—in which all bidders on the supply side received the marginal, market-clearing price—might have borne some responsibility for this phenomenon and the plausible consequent case for paying each supplier only what it bid on its offerings:

> This kind of perceived behavior has lent plausibility to the proposal to substitute pay-as-bid for uniform market-clearing prices. It is only the prospect, under the present system, of receiving on all their sales the benefit of the increase in market price caused by withholding some portion of their capacity that large generators can expect to profit from that practice—so goes the reasoning: under pay-as-bid, in contrast, such generators would have to bid the estimated monopolistically elevated price on all their proffered sales in order to reap those gains, at the immensely increased risk that some or all of those higher bids will prove to have been excessive and therefore be rejected, with a consequent loss of the entire difference between their actual marginal costs and the ultimate market price. The proposed change in the pricing method would, by this reasoning, therefore dramatically alter the balance of risks and potential gains.
>
> Just as the naïve expectation that a shift to pay-as-bid will produce a dramatic reduction in the average prices consumers pay ignores the certainty that generators will radically alter their bidding practices to frustrate achievement of that result, however, so here, the expectation that it would discourage monopolistic withholding fails to take into account the ways in which bidders will respond by changing their bidding behavior correspondingly. If and to the extent that monopolistic withholding has occurred in the past, bidders would henceforward, under pay-as-bid, attempt to predict the consequent behavior of the market prices in their several bids and, to the extent they succeed, the anticipated benefits of the change for consumers will prove to have been illusory.

attention by John Kwoka (Kwoka 2001)—have readily identified California's particular mistake as deregulating only partially. Clearly, the mandated reduction and freeze of retail prices, accompanied by deregulation of wholesale prices, has proved catastrophic, initially to the distribution companies but then, inevitably, to consumers and taxpayers called upon to bail them out. Yet the freeze must have seemed at the time a logical part of a pragmatic resolution of the essentially political problem of providing the companies reasonable assurance of recovery of their stranded costs (to which I, as a former practitioner of prudent investment regulation, felt constrained to acknowledge their entitlement), which was apparently their almost exclusive concern, in exchange for giving ratepayers comparable transitional protection. Indeed for the first couple of years it worked unexpectedly well: wholesale markets were weak and the companies were able to use the expanded margins between acquisition cost and their frozen rates to recover their strandable costs ahead of schedule. How, nevertheless, their managements could have blithely, in effect, bet their companies on the expectation of continued soft wholesale markets, without hedging, is itself difficult to comprehend—at least in hindsight. Moreover, as several have observed, the unexpectedly high prices at which they succeeded in selling off their generating plants might or should have suggested that those expectations were probably wrong.

It may well be that setting aside this particular aspect of the deregulatory design in California—and the deplorable populism and misguided consumerism of promising the public the benefits of unleashed competition while shielding them from its risks[28]—is tantamount to asking, ''Other than that, how did you enjoy the play, Mrs. Lincoln?'' It seems to me necessary, however, to ask whether other aspects of the design may also have been at fault.

4.2.2. Vertical Divestiture

The apparent success of the mandatory divestiture of AT's putatively naturally monopolistic local service from its potentially competitive long distance service[29] has clearly encouraged legislators and regulators to prescribe similar protections in other industries in process of deregulation: observe my criticism, above, of the FCC's infatuation with fully separated subsidiaries and the reluctance of the Department of Justice to give up the interLATA prohibition as the RBOCs seek its lifting (Kahn 1998, 66–69).

In the circumstances, it has seemed natural for several of the state commissions

---

28 Severin Borenstein informed me that the statutorily-mandated 10% reduction in residential rates, which went into effect on January 1, 1998, was not reversed until three years later, and while increases averaging 18% went into effect in May, about one half of all households were to see no increase.

29 But see Crandall (2001). See also the citation of evidence from Canada and Japan that competition in long-distance service had progressed more rapidly in those two countries than in the United States, even though neither of them had required vertical separation of long-distance and local service, in Grieve and Levin (1997); and, similarly, by Spiller and Cardilli that facilities-based *local* competition had progressed at a healthy pace in the smaller countries they examined (Australia, Chile, Guatemala and New Zealand), even though none of these countries has the extensive unbundling requirements for an indefinite duration that prevail in the United States or has prevented incumbents from vertically integrating (1997, 127–138).

similarly to require severing of the ties between electric generating companies and both their transmission and local distribution networks—in order to ensure equal, non-discriminatory access of competitive generators[30] and marketers to those essential facilities. Both California and New York have taken the divestiture route, whether mandatorily or by the offer of sufficient inducements—in part, additionally, in order to secure a market estimate of stranded costs.

I have wondered whether this effectively mandatory divestiture may have contributed to California's problems by denying it one of the main benefits of vertical integration itself—the incentives it provides for undertaking investments at one horizontal level necessary for the success of operations at another. Continued ownership of generation capable of satisfying a larger proportion of their needs would of course have protected the major distribution companies that had accepted the obligation to freeze their retail rates. But whether it would have produced more generation capacity sooner is questionable: there seems to be more than sufficient interest of independent generators in building new plants, as evidenced by the huge number of pending applications for licenses: the assumption that competitive entry into generation is or should be a sufficient protector of consumers—in the long run!—seems to have been borne out.

Whether the same complacency is justified at the transmission level is a question crying out for an answer: when I read in the *Wall Street Journal*—in the news section, mind you, not its editorial pages—that there is inadequate incentive in California to construct new transmission facilities (perhaps the reality would be better conveyed by substituting for ''construct'' ''run the gauntlet of obtaining the necessary permissions to construct''), I cannot but raise the question of whether the almost automatic recourse of some regulators to MFJ-like remedies—apparently Pennsylvania and Illinois have been outstanding, successful exceptions[31]—was wise, especially in an industry in which the vertical interdependencies are so extraordinarily close. At the very least, the California deregulators—in contrast with those in New York, Illinois and Pennsylvania—cannot escape blame for the compulsive tidiness and high propensity to micro-meddle that led them to require utility distribution companies to purchase all their power in the spot market—surely an excessively regulatory accompaniment to what purported to be a process of deregulation. As is widely recognized, distribution companies that adopted the commitment to freeze their rates should, in any sensibly devised system, have had available to them the alternative quasi-integration option of entering into long-term supply contracts, which would similarly have insulated them from what proved after the fact to have been a gigantic risk.

And of course the enforced artificial concentration of purchases and sales on the spot or quasi-spot markets was directly responsible for the immediate explosive effect of the suddenly emerging shortages at times of peak demand on not just the profits of the companies whose retail prices were still frozen—profits almost instantaneously more than wiped out—but also on the rates of the customers unprotected by such freezes, and

30 See Joskow (2000), 128–130.

31 See the references to the healthy progress of Illinois and Pennsylvania—neither of which has required divestiture—in developing new generation capacity in Rowe et al. (2001, 12, 15–16).

inevitably, within a very short time, on the ''protected'' customers as well. It also created the opportunity to which I have already referred for generators with individual generating units large enough to make the difference between adequate and inadequate supplies—in markets characterized by highly inelastic supply and short-term demand—dramatically to lever the market price upward by withholding that output.

4.2.3. Misguided Populism

One widely cited explanation for the prohibition of utility distribution companies acquiring significant proportions of their supplies with long-term contracts was the fear I have seen expressed that they would ''muck it up'' [sic]. That would seem to be a reasonable concern so long as it was contemplated that those companies would continue to operate under the old cost-plus regulatory system. But surely the price freeze to which they were subjected during the transition and the effective competition to which it was anticipated they would be subject thereafter were the proper preventives, remedies or punishment for irrational contracting on their part. Moreover, this concern surely came with ill grace from regulators and strong advocates of independent, non-utility generation who had themselves ''mucked it up'' by forcing on the utility companies purchase prices in their contracts with qualifying facilities under PURPA wildly above avoided costs—from the perspective of the last two decades of the century but no longer!

It appears also the utility distribution companies may have had some legitimate basis in opportunistic behavior of their regulators in the past for a reluctance to avail themselves of the opportunity to hedge their obligation to serve with long-term purchase contracts. As our Committee described that experience:

> The IOUs were able to purchase via forward contracts through the CalPX—but only through the PX—after it first made block forward contracts available in July of 1999. Although there are regulatory limits on the amount of forward positions that each IOU may take, they have not reached those limits, in availing themselves of this opportunity. They may have been discouraged from doing so by their past experience with forward purchases in California, in both gas and electric, under which regulators forced them to absorb any losses stemming from the contract prices exceeding wholesale market prices while not being permitted to reap the benefits, when the contractual prices were lower.[32]

Another explanation of the prohibition of long-term contracts proffered to our Committee was the fear expressed by consumer organizations that such contracts would enable large, business customers to ''skim the cream'' of the benefits available from competition, at the expense of small users. Of course, even if this concern were legitimate, it would justify prohibiting the distribution companies entering into long-term sales, not purchase contracts. The apparent purpose of the latter prohibition was, instead, to preserve equal competitive opportunities for independent generators by preventing the utility companies that continued to own generating plants from favoring those affiliates in obtaining their supplies. Since, however, the independent generators—including purchasers of divested

32 Kahn et al. (2001, original report, footnote 15).

utility plants—were subject to no such requirement to sell exclusively through the organized exchanges, the effort in this meddlesome way to prevent ''cream skimming'' was probably futile, as well as—by denying distribution companies the ability to hedge their risks—injurious.

One final manifestation of misguided, albeit understandable, populism to which I have already referred—the insistence on conferring on consumers the anticipated benefits of deregulation while also protecting them from possible injury from exposure to the vicissitudes of competitive markets—was by retaining the obligation of the utility distribution companies to serve as the suppliers of last resort at regulated prices. The aforementioned Mr. Watts points out that distribution companies cannot enter into long-term purchase contracts so long as they remain subject to the supplier of last resort obligation, because it involves playing a game of heads-we-lose, tails-somebody else-wins: if the terms of the contract prove to be unfavorable—that is, if open market prices fall below the contractual level—they lose all their retail customers and are stuck with the difference between the contractual price and open market value of the power; if and when instead the contract proves favorable, retail customers either remain with them or return to them, in which event they merely break even, at best. If the facts are as he states them, one protective provision eliminated the possibility of another.

4.3. Rate Caps?

The reported increase in California's total electricity bill from $7.4 billion in 1999 to more than $27 billion in 2000 and, according to recent projections (subject to all my preceding admonitions about forecasts in this industry!), $50 billion or more this year, have understandably generated demands for the imposition of ceilings on wholesale rates—specifically, since the traffic is clearly interstate, for the Federal Energy Regulatory Commission to exercise its authority to establish ''just and reasonable'' rates.[33] This appeal—my advocacy of which some commentators have characterized as apostasy—elicited the immediate response, from the President on downward, demonstrating a firm command of the first week of Economics 101:[34] that it would not only do nothing to solve the fundamental imbalance of supply and demand but, particularly by discouraging the expansion of capacity, would exacerbate the problem. I can respond most efficiently in the form of a few pertinent facts and related propositions from the second week.

Both the supply and the demand for electricity, at times when the latter presses hard on capacity, are extraordinarily inelastic. Within the limits of installed capacity, supply is highly responsive to the wide fluctuations of demand from one moment, day or season to the next, as generating stations are automatically brought on line in rank order of their operating costs. When that limit is reached, however, it takes years to add to capacity. True, the elasticity response of demand in California has been prevented by the

33 See the letter initiated by Frank Wolak and signed by nine other economists, including me, calling upon FERC to exercise that authority (May 25, 2001).

34 See the scarcely less simplistic letter to the President of May 29, 2001 by Annelise Anderson, Martin Anderson, Michael J. Boskin, Steve Entin, Milton Friedman and James Gwartney and the much more sensible ''Dim Bulbs'' (Hazlett 2001).

ridiculously extended freeze in charges to residential customers, followed by grudgingly inadequate increases. Since, however, the extreme shortages, producing 10- and 20-fold increases in wholesale prices, have taken the form of extreme spikes at particular times and places,[35] and the overwhelming majority of customers do not have meters permitting them to be charged correspondingly fluctuating prices, the result has been very costly blackouts. It is a truism that blackouts have occurred because most retail rates have not been free to increase to whatever extent necessary to prevent them: what is frightening to contemplate is the extent of that ''extent,'' in the absence of real-time metering.

As against the minimal contribution of such exploding prices to an improved balance of supply and demand at such times[36] must surely be weighed not merely the income distributional consequences but the adverse macroeconomic consequences of generators extracting tens, indeed scores, of additional billions of dollars a year from consumers, businesses and government—much like the contribution to stagflation of the three-fold increases in the prices of crude oil nationally in 1973 and again in 1979–1980.

The spectacular historical instances of price controls doing more harm than good by interfering with the expansion of supply cited routinely by opponents of the proposed FERC initiative—notably on crude oil and natural gas in the 1960s and 1970s—have been ones in which regulation held prices below short- and long-term marginal supply costs (in the case of crude oil, the price of imports). The current electricity price cap proposals with which I am familiar are all intended to incorporate flexibility sufficient to accommodate incremental supplies at high short-run marginal costs, which are readily ascertainable and easily adduced by any generator in support of an exemption; and to ensure the presence of optimal levels of reserve capacity as well.[37]

There seems good reason to believe that the explosion of wholesale prices was not a phenomenon of pure competition alone, but reflected the not-necessarily-collusive or antitrust-law-violative withholdings of capacity at peak times, in order to lever up the market-clearing prices, in a process I have already described. In such circumstances, the rote repeaters of the litany that price controls always reduce supply might be interested in discovering, ceilings may actually result in expansions of offerings: with price caps there is no benefit in withholding supply, only sacrificed profits.

Interference with the fundamentally required correctives—of expanded capacity, on the

---

35 In response to a remonstrance by an independent generator, I am attempting to assess the possibility that in a deregulated system, such price spikes become the necessary mechanism for recovering the capital costs of their peaking units.

36 See note 35, immediately preceding.

37 In a preceding draft, I attempted to rebut the widespread assertions that ceilings inevitably conflict with long-run expansions of supply by citing the experience of the electric industry itself during the entire three-quarter-century, 1920–1995, when, if anything, cost-based regulation, as traditionally practiced, encouraged the very excess capacity that seemed during the past decade to promise such enormous benefits to consumers if rates were deregulated. Upon further reflection I see that this experience could be regarded as irrelevant in a world that had abandoned rate base/rate of return regulation of franchised monopolies, with its implicit guaranteed recovery of costs not imprudently incurred. My colleagues at NERA are helping me evaluate the efficiency of alternative arrangements for ensuring the recovery of the capital costs of capacity required to achieve optimal loss of load probability under the new restructuring arrangements. See also Hobbs et al. (2001).

one side, and intensified conservation, on the other—would of course be severely counterproductive. But where those supply-and-demand responses inevitably take time—to cite the most relevant example, at least a couple of years before additional generating capacity is likely to come on the market—any discouragements can readily be prevented by making the price caps (a) designedly temporary, automatically sunsetting within, say, two to three years, and/or (b) inapplicable to new capacity coming on line. Clearly, however, it is essential that they incorporate some such assurance.

5. Conclusion

The experience of the airlines, telecommunications and electric utility industries over the last 30 years amply documents the superiority of competition over comprehensive regulation. It also demonstrates that successful deregulation rarely consists in total laissez-faire. On the other hand, it amply documents the Tar-Baby effect: as I characterized it some 23 years ago,

> Each time the [regulatory] dike springs a leak, plug it with one of your fingers; just as a dynamic industry will perpetually find ways of opening new holes in the dike, so an ingenious regulator will never run out of regulatory fingers.[38]

In the case of the airlines, the greatest failures and sources of consumer complaint have been the increasing congestion and delays at particular times and places and the extreme differentiation of fares—in particular, the very sharp increase in full fares. The proper remedy for the former is neither governmental nor collaborative airline rationing of scarce air and airport space or promulgation of ''consumers' bills of rights'' but restructuring of the agencies responsible for providing the requisite infrastructure and pricing it correctly; and for the latter, not prescription of maximum fare differentials but intensification of efforts to provide fair competitive opportunities for challengers.

Something approximating total deregulation is clearly infeasible in telecommunications and electric power, essentially for the reasons that led them to be identified as regulated natural monopolies in the first place. At the same time, telecommunications provides the most extreme exemplification of the Tar-Baby effect, because of the demonstrated tendency of regulators to assume responsibility not merely for establishing the necessary conditions for efficient competition but for guaranteeing the desired results. The expansiveness of the FCC's definition of the elements of incumbent telephone company networks that it requires them to make available to entrants and its presumption in prescribing charges equivalent to the results that in its (mistaken) judgment would flow from efficient competition not only flatly discourage facilities-based competition but, as applied to new and extraordinarily expensive facilities, requiring costly and risky investments in latest technologies, conflict directly with the Schumpeterian requirements for dynamic competition.

---

38 Kahn (1978b, pp. 17–18).

Even more than in telecommunications, it is no accident that the proper characterization of the required regulatory reform in the electric power industry is characterized as ''restructuring'' rather than mere deregulation. Fully as much as the FCC's practices, however, the California experience illustrates the evils of deregulatory and regulatory opportunism, of excessively detailed design and prescription and misguided populism—promising consumers the anticipated benefits of competition while sheltering them from its risks. The central task is to design the inescapable residual regulation—centrally, of transmission—in such a way as, by the use of price, to ensure the same kind of efficient synchronization of investments and operations in generation, on the one side, and transmission, on the other, as were previously achieved within vertically integrated regulated suppliers; and, whether by explicit orders or organization of associated markets for capacity, ensuring reserve generating and transmission capacity sufficient for optimal system-wide reliability.

The only other prescription is that regulators be required, in keeping with current fashionable trends in national education policy, regularly to pass tests of their comprehension of Uncle Remus' ''The Wonderful Tar Baby.''[39]

References

Baumol, W. J., J. C. Panzar, and R. D. Willig. 1988. *Contestable Markets and the Theory of Industry Structure*. San Diego: Harcourt Brace and Jovanovich.

Baumol, W. J., and J. G. Sidak. 1994. ''The Pricing of Inputs Sold to Competitors.'' *Yale Journal on Regulation* 11(1): 171–202.

Borenstein, S., and J. Bushnell. 1999. ''An Empirical Analysis of Market Power in a Deregulated California Electricity Market.'' *Journal of Industrial Economics* 47(3): 285–323.

Borenstein, S., J. Bushnell, and F. Wolak. 2000. ''Diagnosing Market Power in California's Deregulated Wholesale Electricity Market.'' POWER Working Paper PWP-064, University of California Energy Institute.

Crandall, R. 2001. ''The Failure of Structural Remedies in Sherman Act Monopolization Cases.'' AEI-Brookings Joint Center for Regulatory Studies Working Paper 01–05.

Demsetz, H. 1974. ''Two Systems of Belief About Monopoly.'' In *Industrial Concentration: The New Learning*, edited by H. J. Goldschmid, H. M. Mann, and J. F. Weston. Boston: Little, Brown and Company, 164–184.

Faulhaber, G. R. 1987. *Telecommunications in Turmoil*, rev. edition. Cambridge: Ballinger.

Gray, H. M. 1942. ''The Passing of the Public Utility Concept.'' Reprinted from *The Journal of Land and Public Utility Economics*, XVI, February 1940. American Economic Association, *Readings in the Social Control of Industry*. Philadelphia: Blakiston: 280–303.

Grieve, W., and S. L. Levin. 1997. ''Telecom Competition in Canada and the U.S.: The Tortoise and the Hare.'' *Selected Papers from the 25th Annual Telecommunications Policy Research Conference*, Alexandria, VA, September 27–29.

Harris, J. C. 1881. ''The Wonderful Tar Baby.'' *Uncle Remus, His Songs & His Sayings: The Folk-Lore of the Old Plantation*. New York: Appleton and Company.

Hazlett, T. W. 2001. ''Dim Bulbs.'' AEI-Brookings Joint Center, *Policy Matters* 01–03, January.

Hobbs, B. F., J. Iñón, and S. E. Stoft. 2001. ''Installed Capacity Requirements and Price Caps: Oil on the Water, or Fuel on the Fire?'' *The Electricity Journal* (July): 23–34.

Hogan, W. W. 2002. ''Electricity Market Restructuring: Reforms of Reforms,'' Presented at CRRI's 20th Annual Eastern Conference, May 25, 2001, *Journal of Regulatory Economics* 21(1).

39 McKie (1970, p. 9) first used the term ''tar baby effect'' from the original story (Harris 1881).

Joskow, P. L. 2000. ''Deregulation and Regulatory Reform in the U.S. Electric Power Sector.'' In *Deregulation of Network Industries: The Next Steps*, edited by S. Pelzman and C. Winston. Washington, DC: AEI-Brookings Joint Center for Regulatory Studies, 113–188.

Joskow, P. L., and E. Kahn. 2001. ''A Quantitative Analysis of Pricing Behavior in California's Wholesale Electricity Market during Summer 2000.'' National Bureau of Economic Research Working Paper #8157.

Kahn, A. E. 1970 and 1971. *The Economics of Regulation* (2 volumes). New York: John Wiley & Sons. Reprinted by MIT Press, 1988.

Kahn, A. E. 1978a. ''Route Awards and Airline Scheduling Practices.'' Remarks before the Federal Aviation Agency (March 22).

Kahn, A. E. 1978b. Presentation before a Symposium on the Changing Environment of International Air Commerce, Georgetown University, Washington, D.C., May 4.

Kahn, A. E. 1988. ''Airline Deregulation: 'I Would Do It Again'.'' *The Washington Post* (September 25): C6.

Kahn, A. E. 1993. ''Competitive Consequences of Hub Dominance: A Case Study.'' *Review of Industrial Organization* 8: 381–405.

Kahn, A. E. 1998. *Letting Go: Deregulating the Process of Deregulation*. Michigan State University Institute of Public Utilities.

Kahn, A. E. 2001. *Whom the Gods Would Destroy, or How Not to Deregulate*. Washington, D.C.: AEI-Brookings Joint Center for Regulatory Studies.

Kahn, A. E., and W. B. Shew. 1987. ''Current Issues in Telecommunications Regulation Pricing.'' *Yale Journal on Regulation* 4(2): 191–256.

Kahn, A. E., and W. E. Taylor. 1994. ''The Pricing of Inputs Sold to Competitors: A Comment.'' *Yale Journal on Regulation* 11(1): 225–240.

Kahn, A. E., T. J. Tardiff, and D. L. Weisman. 1999. ''The Telecommunications Act at Three Years: An Economic Evaluation of Its Implementation by the Federal Communications Commission.'' *Information Economics and Policy* (December): 319–365.

Kahn, A. E., P. C. Cramton, R. H. Porter, and R. D. Tabors. 2001. ''Pricing in the California Power Exchange Electricity Market: Should California Switch from Uniform Pricing to Pay-as-Bid Pricing?'' A study commissioned by the California Power Exchange, January 23; published in amended form as ''Uniform Pricing or Pay-as-Bid Pricing: A Dilemma for California and Beyond'' in *The Electricity Journal* 14 (July): 70–79.

Kaserman, D., and J. T. Mayo. 1991. ''The Measurement of Vertical Economies and the Efficiency Structure of the Electric Utility Industry.'' *Journal of Industrial Economics* 39: 483–502.

Kuttner, R. 1988. ''High Costs of Flying.'' *The Boston Globe* (September 16).

Kwoka, J. 2001. ''California: The Perfect Electrical Storm.'' AAI Column in *FTC: Watch* (February 14).

Levine, M. E. 1969. ''Landing Fees and the Airport Congestion Problem.'' *Journal of Law and Economics* XII (April): 79–108.

Levine, M. E. 1987. ''Airline Competition in Deregulated Markets: Theory, Firm Strategy and Public Policy.'' *Yale Journal on Regulation* 4(2): 392-494.

MCI v. FCC, CCA, DC Circuit, July 28, 1977, *Utility Law Reports*, Federal, Paragraphs 12,000–12,024.

New York Times. (2001) ''Canada Plays Tough on Competitors.'' (April 22, V:3).

McKie, J. W. 1970. ''Regulation and the Free Market: the Problem of Boundaries.'' *The Bell Journal of Economics and Management Science* 1(1): 6–26.

Noll, R. G., M. J. Peck, and J. J. McGowan. 1973. *Economic Aspects of Television Regulation*. Washington: Brookings Institution.

Robyn, D. 2001. ''Congress' Plan to Reduce Flight Delays is Not Airworthy.'' AEI-Brookings Joint Center for Regulatory Studies, Policy Matters 01–15 (May).

Rosen, R. A., F. Sverrisson, and J. Stutz. 2000. *Can Electric Utility Restructuring Meet the Challenges It Has Created?* Boston: Tellus Institute.

Rowe, J. W., P. Thornton, and J. B. Szczypinski. 2001. ''Competition Without Chaos.'' AEI-Brookings Joint Center for Regulatory Studies, Working Paper 01–07, June.

Spiller, P. T., and C. G. Cardilli. 1997. ''The Frontier of Telecommunications Deregulation: Small Countries Leading the Pack.'' *The Journal of Economic Perspectives* 11: 127–138.

Transportation Research Board/National Research Council. 1999. ''Entry and Competition in the U.S. Airline Industry, Issues and Opportunities.'' Special Report 255. Washington, D.C.

U.S. District Court for the District of Kansas. 2001. *Memorandum and Order, 2000, U.S.A. v. AMR Corporation et al.* Internet version (April): 49–51.

Watts, P. C. (pseudonym of Michael Rothkopf). 2001. ''Heresy? The Case Against Deregulation of Electricity Generation.'' *The Electricity Journal* (May): 19–24.

Revised November 7, 2000

ANTITRUST IN THE NEW ECONOMY

Richard A. Posner[1]

Concern has been expressed recently that U.S. antitrust law may not be well suited to regulating the "new economy." Doctrines developed to deal with competition and monopoly in smokestack industries is not well adapted, it is argued, to dealing with the dynamic economy of the twenty-first century. What I shall argue is that there is indeed a problem with the application of antitrust law to the new economy, but that it is not a doctrinal problem; antitrust doctrine is supple enough, and its commitment to economic rationality strong enough, to take in stride the competitive issues presented by the new economy. The real problem lies on the institutional side: the enforcement agencies and the courts do not have adequate technical resources, and do not move fast enough, to cope effectively with a very complex business sector that changes very rapidly. This problem will be extremely difficult to solve; indeed, I cannot even glimpse a complete solution, although I will suggest some partial solutions.

I.

I shall use the term the "new economy" to denote three distinct though related industries. The first is the manufacture of computer software. The second consists of the Internet-based businesses (Internet access providers, Internet service providers, Internet content providers), such as AOL and Amazon. And the third consists of communications services and equipment designed to support the

---

[1] Judge, U.S. Court of Appeals for the Seventh Circuit; Senior Lecturer, University of Chicago Law School. Previous versions of this paper were given as talks in New York City on September 14, 2000, at a conference on antitrust sponsored by the American Bar Association and the American Law Institute, and the next day at a meeting of the Council of Foreign Relations; at DePaul University in Chicago on September 21, 2000, as a lecture in the Distinguished Lecture Series in Finance; as a presentation at the Law and Economics Workshop of the University of Chicago Law School on October 2, 2000; and as a lecture at the University of Hamburg on November 6, 2000. I thank Dennis Carlton, Frank Easterbrook, Eric Posner, William Landes, and participants at the various sessions at which the paper was given for their very helpful comments.

first two markets. There are other candidates for inclusion in the new economy, but these three will do for my purposes.

These industries differ markedly from most of the industries in which modern antitrust doctrine emerged, and particularly from industries that manufacture traditional physical goods such as steel, automobiles, pipe, wire, aluminum, railroad cars, roadbuilding materials, and cigarettes. The traditional industries are characterized by multiplant and multifirm production (indicating that economies of scale are limited at both the plant level and the firm level, or in other words that average total costs are rising at relatively modest output levels), stable markets, heavy capital investment, modest rates of innovation, and slow and infrequent entry and exit. The new-economy industries that I'll be discussing tend to lack these features. They are characterized instead by falling average costs (on a product, not firm, basis) over a broad range of output, modest capital requirements relative to what is available for new enterprises from the modern capital market, very high rates of innovation, quick and frequent entry and exit, and economies of scale in consumption (also known as "network externalities"), the realization of which may require either monopoly or interfirm cooperation in standards setting. And while vertical integration is a common feature of the old economy, it tends to be even more common in the new one, precipitating an unusually large number of firms into customer or supplier relations with other firms that are also its competitors.

A key to understanding these industries is that, with the partial exception of communications equipment, their principal output is intellectual property, namely computer code, rather than physical goods. (Although the intellectual property may be shipped on a disk, this is incidental, and anyway software is increasingly shipped to the purchaser over the Internet.) This is obviously the case for computer software but is also true to a large extent of the Internet-based businesses, even those that sell physical products such as groceries and books. Their ability to take and fill orders and carry out the other operations (such as marketing, billing, handling returns, responding to customer questions and complaints, and allaying the customer's privacy and security concerns) required to give their customers whatever goods or services the business provides is a function to a large extent both of the sophistication of the business's computer software and of its trademarks and copyrights.

Intellectual property is characterized by heavy fixed costs relative to marginal costs. It is often very expensive to create, but once it

is created the cost of making additional copies is low, dramatically so in the case of software, where it is only a slight overstatement to speak of marginal cost as zero.[2] Without legal protection, the creator of intellectual property may be unable to recoup his investment, because competitors can free ride on it; and so legal protection can expand output rather than, as in the usual case of monopoly, reduce it. At the same time, legal protection of intellectual property may operate to deflect consumers to more costly substitutes. The owner of the patent or copyright will charge a positive price for copying, even though the marginal cost may be zero; the positive price will deflect some consumers to substitutes. To prevent these defections would require perfect price discrimination, which is infeasible because (administrative costs to one side) it would require the seller to have complete information about the elasticity of the demand for his product by all his customers and potential customers.

The patent and copyright laws try to strike the output-maximizing balance by giving the creator of intellectual property some but not complete protection from competition. A copyright confers protection for a longer period than a patent does because, traditionally, less was protected—just the form in which the composer, painter, writer, etc. had chosen to express his ideas, and just copying, not independent discovery. The extension of copyright to software has been controversial. Many observers believe that it confers excessive protection, in the sense either of restricting rather than expanding the amount of intellectual property or of attracting excessive resources into the production of that property relative to investment in markets that do not yield a monopoly return (or both), because of difficulties in inventing around an innovative code. Even worse, it is argued, the methods of distributing software often enable the creator to obtain by contract even more protection than copyright law gives him. Copyright gives the holder a property right in his intellectual property even when it is in the hands of a person with whom he has no contract, such as the purchaser of a copyrighted book from a bookseller. To the extent that the creator of software contracts directly with the ultimate purchaser, he can impose via contract more restrictions than the copyright law would allow him to do in the absence of contract;[3] for example, he can forbid the pur-

---

[2] But still an overstatement, because there are selling and servicing costs associated with each sale or rental of software.

[3] See Mark A. Lemley, "Beyond Preemption: The Law and Policy of Intellectual Property Licensing," 87 *California Law Review* 111, 124–134 (1999).

chaser to make an extra copy for his own use, as copyright law permits. It is true that a contract is unlikely to have the same duration as copyright protection, but length of protection is academic in the case of software, which becomes obsolete long before the copyright on it expires. When discrimination is less than perfect, the effect on output is indeterminate, as I'll explain shortly.

The possibility that the combination of copyright and contract gives software manufacturers too much monopoly power in the economic sense, that is, either causes a lessening rather than an increase in the output of the intellectual property in question or distorts the allocation of resources between industries that produce intellectual property and industries that do not, or has both bad effects, creates a natural concern with any further practice or circumstance that might increase the manufacturer's power over the price of his software. It is at this point that another feature of the new economy that I mentioned, economies of scale in consumption, becomes troublesome. Economies of scale in manufacture are familiar; up to a point, the longer the production run, the lower average cost is. Economies of scale in consumption refer to the situation in which the larger the firm's output is (up to some point), the more valuable that output is to its customers. The traditional example was the telephone. Telephone service is worthless if there is only one subscriber; he has no one to talk to. The more subscribers, the more valuable the service is to each one, or at least to many of them. Interactive services, such as email and online auctions, are similar. Likewise the sharing of computer programs, as where two or more academics collaborate on writing a scholarly article by means of word-processing and spreadsheet programs. Literal networking or sharing to one side, computer programs tend to be more valuable the more people use them because training, support by IT personnel, and standardization of equipment and procedures are facilitated. It is the same reason that the typewriter keyboard is standardized. We might call this "figurative" networking, by analogy to the distinction between literal and figurative queuing (for example, standing in line to buy a ticket versus waiting for one's case to be called for trial).

These economies of consumption presuppose uniformity rather than a common source. The international telephone system is a single network, but its components are owned by a vast number of separate firms and individuals. The components have, however, been standardized to assure interoperability, in the same way that the gauge of the railroad track has been standardized throughout the

nation. A firm that manufactures one of the essential components of a network (a term I am using broadly, to denote any situation in which there are economies of scale in consumption) would prefer to be the exclusive source of that component rather than be required to disclose the information that would enable competitors to duplicate it. If the component is subject to intellectual-property protection through patent, copyright, or contract, or can be held as a trade secret, then the requisite uniformity may be more readily achievable by monopoly provision than by standardization.

Networks are not valuable to the consumer in themselves; they are conduits for the services that the consumer values. This is one point at which vertical integration enters the new economy. An operating system is a platform for software applications, and so the writer of operating-system software may decide to write software applications to ride on it, in much the same way that AT&T used to manufacture the terminal equipment attached to its telephone lines. Modern operating systems are themselves composites of separate programs which may be provided by separate companies or by one company; in the latter case there is an analogy to AT&T s practice of manufacturing the switching equipment for its telephone system as well as the telephone lines themselves. Firms that provide dial-up connections and other facilities for accessing and browsing the World Wide Web can integrate forward into the provision of Web-based services such as shopping and video.

The features of the new economy that I have been describing, all but the last (vertical integration, of which more later), tug it toward monopoly yet also, oddly, toward competition. The paradox is dissolved by remembering that competition to *obtain* a monopoly is an important form of competition. The more protection from competition the firm that succeeds in obtaining a monopoly will enjoy, the more competition there will be to become that monopolist; and provided that the only feasible or permitted means of obtaining the monopoly are socially productive, this competition may be wholly desirable. A firm that will have the protection both of intellectual-property law and of economies of scale in consumption if it is the first to come up with an essential component of a new-economy product or service will have a lucrative monopoly, and this prospect should accelerate the rate of innovation, in just the same way that, other things being equal, the more valuable a horde of buried treasure is, the more rapidly it will be recovered. What is more, the successful monopolist is likely to be a firm that initially charges a very low price for the

new product that it has created. Think back to the telephone. Since every new subscriber increases the value of the service to the existing subscribers, a telephone company has an incentive to provide price inducements to new subscribers, as the money it will lose on them may be more than made up for by the higher price that existing subscribers will pay for access to a larger network. This is especially likely if the network will be a natural monopoly, in the sense that no competitor would find it feasible to duplicate it—then the faster the network reaches maturity the longer the monopolist will be protected from challenges to his monopoly. The prospect of a network monopoly should thus induce not only a high rate of innovation but also a low-price strategy that induces early joining and compensates the early joiners for the fact that eventually the network entrepreneur may be able to charge a monopoly price.

I emphasize "may be able" in the preceding sentence. Traditional networks such as the telephone system and the railroads required enormous capital investments and were therefore difficult to duplicate. If you owned such a network, or an important part of it, you had a pretty secure monopoly. The less capital investment the creation of a substitute network involves, the less secure the network monopolist's monopoly is. Because of the extraordinary pace of innovation not only in computer software but also in communications technology, the extraordinary amount of capital that is available worldwide for investment in new enterprises, and the rapidity with which new networks that are primarily electronic can be put into service, the networks that have emerged in the new economy do not seem particularly secure against competition. We have seen all manner of firms rise and fall in this industry—falling sometimes from what had seemed a secure monopoly position. The gale of creative destruction that Schumpeter described, in which a sequence of temporary monopolies operates to maximize innovation that confers social benefits far in excess of the social costs of the short-lived monopoly prices that the process also gives rise to, may be the reality of the new economy.

The feasibility of challenging an existing network monopolist is critical, however. Even if the only way to become a network monopolist in the new economy is to be the first to come up with a new technology that benefits consumers, the existence of the monopoly may discourage subsequent technological innovation by other firms. If network externalities are large, they may give the monopolist a cost advantage that exceeds the benefit of a superior new technology.

This is the issue of "path dependence": an industry may be stuck with an inferior technology because of the cost advantage of the existing network.[4]

Appearances to the contrary notwithstanding, the antitrust laws are not much concerned with monopoly as such.[5] It is not a violation of those laws to acquire a monopoly by lawful means, and those means include innovations protected from competition by the intellectual-property laws. If copyright protection of computer software is too broad (too much intellectual-property protection can reduce output, and, for that matter, too much innovation can be inefficient too[6]), that is a matter to take up with Congress. Nor is it a violation of antitrust law to charge a monopoly price, or to price discriminate in an effort to maximize monopoly profits. And now that the *Alcoa* doctrine[7] is discredited, it is understood that a monopolist is free to compete, whether against the competitive fringe in its monopoly

---

[4] See, for example, Stanley M. Besen and Joseph Farrell, "Choosing How to Compete: Strategies and Tactics in Standardization," *Journal of Economic Perspectives*, Spring 1994, p. 117; Joseph Farrell and Garth Saloner, "Installed Base and Compatibility: Innovation, Product Preannouncements, and Predation," 76 *American Economic Review* 940 (1986); Michael L. Katz and Carl Shapiro, "Technology Adoption in the Presence of Network Externalities," 94 *Journal of Political Economy* 922 (1986).

[5] This is what makes Eastman Kodak Co. v. Image Technical Services, Inc., 504 U.S. 451 (1992), involving the tying of repair services to the purchase of the product, along with most other tying cases, so questionable, as it was plain that, at worst, Eastman Kodak was merely exploiting its customers, not excluding or discouraging equally or more efficient competitors. Benjamin Klein, "Market Power in Antitrust: Economic Analysis after *Kodak*," 3 *Supreme Court Economic Review* 43 (1994).

[6] A patent confers a lawful monopoly, but the expenditures on obtaining a patent may exceed the social benefits of that lawful monopoly. Imagine a case in which those social benefits, which I'll assume are equal to the royalties that the patentee will obtain, are $10 million, and the optimal path to creating the patentable innovation involves two firms each spending $4 million. The firm that creates it first will obtain the patent. Suppose that after each has spent $3 million, they estimate that it will take another six months to complete their respective patent applications. Since the expenditure of $6 million is a sunk cost, each will spend up to an additional $10 million to complete his application a day before his rival; but the social benefit of the acceleration in the completion day will be negligible. Monopoly in this example draws excessive resources into innovating.

[7] United States v. Aluminum Co. of America, 148 F.2d 416 (2d Cir. 1945) (L. Hand, J.).

market or against potential competitors, as vigorously as a firm in an ordinary competitive market would be, provided it doesn't employ tactics calculated to drive an equally or more efficient firm from the market.[8] The fact that a monopolist buttressed by network externalities may be hard to dislodge even by a firm with a superior technology has no antitrust significance in itself. What is true is that a firm is forbidden to enter into a price-fixing agreement with its competitors, or to acquire a competitor if the acquisition will alter the structure of the market to make it much more conducive to price fixing, but these rules do not require any modification to deal with mergers and price fixing in new-economy industries. Nor is there a flat rule against communication among competitors on matters such as standards setting where consumers may benefit from a degree of interfirm cooperation.

The focus of concern with the application of antitrust law to the new economy is on the methods by which a firm that has a monopoly share of some market in a new-economy industry might seek to ward off new entrants. The lawyers and economists who express this concern are fearful lest a "Chicago school" approach to antitrust deny the possibility that a single firm, without collaborating with competitors or potential competitors (thus inviting application of the rules against price-fixing and large horizontal mergers), can, at least under new-economy conditions though probably more generally as well, prevent efficient challenges to its monopoly. If the Chicago-school approach so understood is law, these critics want it modified to do service in new-economy antitrust cases.

This is a misunderstanding of the Chicago school, at least if I can be considered a member of the school in good standing.[9] The approach is skeptical—but no stronger word would be correct—about the danger to competition that is posed by unilateral firm action,

---

[8] See, for example, United States v. Syufy Enterprises, 903 F.2d 659, 669–669 (9th Cir. 1990); Olympia Equipment Leasing Co. v. Western Union Telegraph Co., 797 F.2d 370, 375–376 (7th Cir. 1986); Foremost Pro Color, Inc. v. Eastman Kodak Co., 703 F.2d 534, 544 (9th Cir. 1983). These cases reflect a series of emphatic post-*Alcoa* statements by the Supreme Court that the antitrust laws protect competition in the sense of efficient business practices rather than in the sense of rivalry per se. See, for example, Broadcast Music, Inc. v. Columbia Broadcasting System, 441 U.S. 1, 19–20 (1979); Reiter v. Sonotone Corp., 442 U.S. 330, 343 (1979).

[9] See Richard A. Posner, *Antitrust Law: An Economic Perspective*, ch. 8 (1976).

unilateral in the special sense that it does not require cooperation with competitors (it usually requires cooperation with customers or suppliers). The approach emphasizes both the difficulty of squashing competition by such means and the danger that heavy-handed antitrust enforcement may suppress a practice that may seem anticompetitive but actually is efficient, or at least neutral, from the broader social standpoint. A classic example is the tying agreement, which used to be thought a means by which a firm having a monopoly of one market (the market for the tying product) could obtain a second monopoly (over the market for the tied product). Overlooked was the fact that the cases involved complementary products (hammers and nails, for example), so that, by definition of complementarity, an attempt to increase the price of one would reduce the demand for the other.[10] Owning both monopolies would produce a net increase in the monopolist's profits only if the second monopoly enabled him to engage in price discrimination more easily, with sales of the tied product being used to monitor the intensity of the consumer's demand for the service (such as hammering nails) that the products jointly produce. Although price discrimination has no general tendency to increase efficiency,[11] banning one form is unlikely to do any good, since there is no general antitrust prohibition against price discrimination.

---

[10] A more fundamental point, however, is that tying *makes* the tied product a complement of the tying one; the buyer cannot get the tying product without the tied product and will therefore pay less for the tying product the higher the price for the tied product. The significance of the fact that the tied product is usually a complement of the tying product independently of the tie is that the quantity demanded of the tied product is a measure of the intensity of the buyer's demand for the tying product.

[11] Perfect price discrimination would bring about the same output as under competition, because no customer willing to pay the seller's marginal cost would be turned away. But perfect price discrimination is infeasible, and imperfect price discrimination can result in a lower or higher output than under competition, or the same output. See Frederick M. Scherer and David Ross, *Market Structure and Industrial Performance* 494–496 (3d ed. 1990); Paul A. Samuelson, *Foundations of Economic Analysis* 42–45 (1947); Joan Robinson, *The Economics of Imperfect Competition* 188–195 (1933). Many economists believe that even crude discrimination is more likely to expand than to reduce output, see, for example id. at 201; Scherer and Ross, *supra*, at 494–496; Peter O. Steiner, Book Review, 44 *University of Chicago Law Review* 873, 882 (1977), but there does not appear to be a firm basis for this belief. See Hal R. Varian, "Price Discrimination," in *Handbook of Industrial Organization*, note 12 above, vol. 2, pp. 597, 629–633.

Or consider exclusive dealing. A manufacturer of some consumer product requires its distributors to agree not to carry any potential competitor's products, thus increasing (he hopes) the cost of distribution to potential competitors. With an important exception noted below for the case where there are economies of scale in distribution, the manufacturer will have to compensate its distributors for agreeing to the restriction, and this will increase its costs, making entry more attractive. If a potential competitor has a promising product, other distributors will be delighted to carry it; if there are no other distributors, new ones will appear. Even if the potential competitor has to do its own distribution, and thus enter on two levels, manufacturing and distribution, the fact that the capital requirements for entry are now greater should not be an obstacle to entry, since there is no shortage of capital for promising new ventures. For these reasons, exclusive dealing is unlikely to be an effective means of forestalling entry. At the same time, it is easy to see how exclusive dealing might promote efficiency by increasing the likelihood that a distributor will use his best efforts to promote the manufacturer's goods. Exclusive dealing may also help a seller of intellectual property to prevent piracy, a serious concern in intellectual-property markets.

But skepticism about unilateral monopolizing actions is not the same as denial. The *Standard Fashion* case[12] has long been used by some members of the Chicago school to illustrate a unilateral action that might well have increased the defendant's monopoly power.[13] The defendant manufactured a line of women's dress patterns (which women used to make their own dresses) that was very popular. Retailers thought it essential to be able to sell the line. The defendant required retailers to agree not to carry competing lines. Competing manufacturers could in principle create their own retail outlets, but who would shop there if the most popular brand could not be found? Competing manufacturers would have to create a line as long and as popular as Standard Fashion's line, and that would be difficult, maybe impossible, to do.

---

[12] Standard Fashion Co. v. Magrane-Houston Co., 258 U.S. 346 (1922).

[13] Aaron Director and Edward H. Levi, "Law and the Future: Trade Regulation," 51 *Northwestern University Law Review* 281, 293 (1956); Richard A. Posner, *Antitrust Law: An Economic Perspective* 202 n. 48 (1976). See generally John E. Lopatka and William H. Page, "Posner's Program for the Antitrust Division: A Twenty-Five Year Perspective," 48 *SMU Law Review* 1713, 1734, 1734 (1995).

What distinguished *Standard Fashion* from a garden-variety exclusive-dealing case was the existence of economies of scale at the distribution level. Consumers didn't want to traipse from store to store. They wanted a full line in each store, so anyone entering the women's clothing business had to provide the full line if it was excluded from stores that carried the dominant firm's line. Restricting its retailers no doubt cost Standard Fashion something. But it is plausible that the cost was less than the increase in its expected monopoly profits from forestalling new entry by compelling any prospective entrant to enter on a full-line basis.

The point is not that the new entrant would have to invest more capital, as in the previous example. The point is that it would have to embark on a riskier undertaking, that of creating not a single successful product but a whole line of such products. It's as if one couldn't make commercial aircraft without making military aircraft as well. In such a case, as in *Standard Fashion* itself, exclusive dealing, while it would increase the defendant's net costs of distribution (if it were indeed adopted for anticompetitive purposes and had no efficiency rationale), might increase the costs (not the capital requirements) of new entrants much more.

The analogy to a new-economy network externality should be plain. The network corresponds to the full-line retail store in *Standard Fashion*. A firm may wish to enter the market by producing one component of the network or one value-added service, but if a competitor by virtue of owning or having an exclusive-dealing contract with the network refuses to cooperate with the firm, the firm won't be able to duplicate the network in order to get distribution of its component or service.

It is true that piecemeal entry is the norm in many industries. A department store carries the products of many producers, many of which do not offer a full line of products. One can imagine a number of designers of women's dress patterns, each specializing in one pattern, and the department store assembling them into a full line to compete with Standard Fashion's full line. The risk of entry to each designer would be minimized. The case was decided in 1922, however, and many towns may not have had department stores. We may be in a similar stage in the development of the new economy, where distribution facilities may be sufficiently limited to create bottlenecks that monopolists can exploit to perpetuate monopoly.

It is important to note that the monopolist would have no incentive to engage in exclusionary conduct unless his monopoly were

fragile, that is, vulnerable to new entry. The more vulnerable it is, however, the less likely it is to endure even if the monopolist does resort to such conduct. This point is missed in Robert Bork's criticism of the decision.[14] He argues that Standard couldn't extract a monopoly price from its dealers twice, once by charging them what the market would bear and then by forcing them to enter into exclusive-dealing contracts. That's true, but what Standard may have been able to do was to increase the duration of its monopoly.

The likelihood that the monopoly profits obtained during the extension period (as we may call the period for which a monopoly is extended by means of exclusionary practices) will exceed the costs of the exclusionary practice to the monopolist is enhanced by the fact that, both in *Standard Fashion* and even more dramatically in parallel new-economy cases, the monopoly is of intellectual property. Intellectual property is characterized as we know by a large difference between total and marginal cost. In the extreme case, which is approximated in some software markets, marginal cost is close to zero, meaning that almost all the revenues earned by a firm that monopolizes the market go directly to the bottom line. This makes it plausible that the profit from extending the monopoly another year or two will exceed the cost of the exclusionary practices required to achieve the extension. To put this differently, there is no reason to think that the cost of an exclusionary practice in such markets will exceed the additional monopoly profits that the practice makes possible, even though the cost may be incurred earlier, which requires that any future profits be discounted to present value before being compared to the cost.

A complicating factor in both *Standard Fashion* and new-economy cases is the utility of exclusive dealing as a means of dealing with piracy. *Standard Fashion* may have been worrying about the piracy of its patterns by competing manufacturers, and attempting to prevent this by denying outlets to the pirates. (This concern was explicit in a later case, *Fashion Originators' Guild of Amer-*

---

[14] Robert H. Bork, *The Antitrust Paradox: A Policy at War with Itself* 305–307 (2d ed. 1993). Bork also argues that Standard Fashion's market share, 40 percent, was too small to enable it to foreclose rivals. But that was its nationwide market share, and the Supreme Court pointed out that in many towns and smaller cities it had a monopoly. It should probably be regarded therefore as a case involving a series of local monopolies rather than a national monopoly.

ica v. FTC.[15]) To anticipate the second part of my discussion, whenever an antitrust court is called on to balance efficiency against monopoly, there is trouble; legal uncertainty, and the likelihood of error, soar.

Exclusive dealing, by the way, is analytically the same as tying. Exclusive dealing ties distribution to manufacture; equivalently, tying is exclusive dealing in the tied product. All you need is sensible law on exclusive dealing to deal sensibly with tying cases. If there are economies of scale in the tied product, a firm that must in order to enter the market for the tying product produce the tied product as well will have higher costs than the monopolist and this will reduce the expected gain to him of entering the market for the tying product. Notice, finally, that exclusive dealing and tying can be accomplished in a variety of different ways, including vertical integration, contract, product design (in the case of tying), and bundling (charging a zero price for the tied product, and so in effect paying the consumer to buy it from the supplier of the tying product). And, speaking of bundling, while documented cases of predatory pricing are rare, it is not always an irrational method of deterring entry. Especially when it takes the form of area price discrimination or "fighting brands," so that the predator does not have to lose more money than the new entrant by lowering prices throughout his market, a monopolist may maximize his profits by "investing" in a reputation for predatory response to threatened entry.

II.

Well, the plaintiff won in *Standard Fashion*, and the decision remains good law and its principle is not limited to exclusive dealing but extends as we have just seen to any business method by which a monopolist can exclude an equally or more efficient competitor; likewise the illegality of predatory pricing is settled antitrust doctrine. I conclude that, as I said at the outset, existing antitrust doctrine is sufficiently supple, and sufficiently informed by economic theory, to cope effectively with the distinctive-seeming antitrust problems that the new economy presents.

What is troublesome is the institutional structure of antitrust enforcement. To begin with, cases in the new economy present unusually difficult questions of fact because of the technical complexity of the products and services produced by new-economy industries.

---

[15] 312 U.S. 457 (1941).

Such questions can be central rather than peripheral to the antitrust issues, as where a plaintiff complains that the defendant has altered a technical protocol or interface to make it more difficult for the plaintiff s product to work with the network, or a defendant contends that disclosure of a protocol would enable its competitors by reverse engineering to deprive it of a valuable trade secret that it cannot feasibly protect by means of copyright or patent law. The second example is an aspect of two broader questions, both also very technical and very difficult—the relative merits of monopoly provision and standardization as methods of optimizing a network, and the adequacy in fact of copyright or patent law relative to trade-secret law to protect investments in software and other new-economy products. Similar questions have arisen in other antitrust network settings, notably in the telephone industry, but they simply are much more difficult for a lay person to understand in the new-economy context.

A further complication is that it is difficult to find truly neutral competent experts to advise the lawyers, judges, and enforcement agencies on technical questions in the new economy. There aren' t that many competent experts, and almost all of them are employed by or have other financial ties to firms involved in or potentially affected by antitrust litigation in this sector. The Antitrust Division does not employ *any* computer scientists or electrical engineers, but is wholly dependent on consultants, as are also, I believe, all the state antitrust offices; and, as I say, it is difficult to find a consultant in the new economy who is both competent and disinterested.

We deal with technical questions in the judiciary not by having judges or jurors who have the requisite technical knowledge or by giving them technical assistants but by having technical experts present evidence which the judge and jury (if it is a jury case) is expected somehow to assimilate. This system does not work as badly as its critics maintain; but the more technical the area of litigation and the fewer experts are disinterested, the worse it is apt to work. Computer science and communications technology are much more difficult areas than the average body of scientific or engineering knowledge that lay judges and jurors are asked to absorb en route to rendering a decision.

One possible measure would be to borrow from arbitration procedure, as follows: In a technically complex case, the judge could direct each party to nominate a technical expert, and the two experts selected in this way would then agree upon a third, a neutral, expert,

whom the judge would appoint as a court-appointed expert[16] in confidence that he would not be misled by a partisan. The problem, as I mentioned earlier, is that there may be very few genuine neutrals in the high-technology sector who have the expertise required for a particular case, and they may be reluctant to become involved in a litigation. Setting that problem aside, a technical committee composed of two party-nominated and one neutral technical expert might assist the judge or special master in the administration of a consent or litigated decree if one were entered in an antitrust case against a new-economy company. This is more needful than in the case of warring economic experts, because the economic principles that bear on antitrust can be explained in intuitive terms; technical principles often cannot be.

The difficult factual questions presented by new-economy antitrust cases are not limited to technical areas, unfortunately. The combination of intellectual property, network externalities, and rapid growth in consumer demand creates difficult questions involving the ascertainment and measurement of monopoly. Suppose a firm has 100 percent of some new-economy product market and charges a price for it that is greatly in excess of marginal cost. Suppose also that the product is one that involves economies of scale in consumption (that is, it is a network market) and the demand for it is growing very rapidly. This is a common new-economy scenario. Does the firm have a monopoly in an economically relevant sense? The ratio of price to marginal cost is meaningless, since often pricing intellectual property at marginal cost is nonremunerative (because marginal cost is lower than average total cost, owing to heavy fixed costs) and leads to bankruptcy. What about the market share? But one characteristic of intellectual property is its durability, and the more durable a product is, the fewer repeat sales the manufacturer will have. (It is different if the firm leases rather than sells its product.) New-economy firms that don't lease get around this problem with frequent upgrades, but the installed base acts as a drag on the price that can be charged for an upgrade, since the prospective customer for the upgrade has the option of sticking with the existing product. If demand is growing rapidly, moreover, the firm in deciding on a price must trade off current against future profits, since a higher price today will slow the growth in demand. And by the assumption that it is a network market, the firm will have a strong in-

[16] As expressly authorized by Fed. R. Evid. 706.

centive to charge a very low price in order to increase usage because that will enable it to charge a much higher price in the future—unless rapidity of innovation makes the future wholly uncertain, in which event it may abandon the hope of a network monopoly and charge a high price now. Still another factor is that a vertically integrated firm competes against itself: a firm that owns a network but also sells value-added services may want to keep the price of network access low to increase sales of those services. This is the other side of the tying coin: holding down the price of the tied product will keep up the demand for the tying product. The result of all these factors may be that a firm has a monopoly market share only because it is not charging a monopoly price.

And while it is possible to argue that monopoly can have other bad effects besides the limitation of output over the competitive level that monopoly pricing brings about, namely a reduction in the rate or distortion in the direction of innovation, economic theory and empirical evidence have yet to generate a consensus on whether monopoly is on balance good or bad for innovation. The possible effect of network externalities in discouraging subsequent innovation (the "path dependence" problem) not only is speculative, but is operative even if the monopolist is passive, yet in that case there would be no arguable antitrust violation.

The peculiarities of new-economy markets that I have been describing are apt to make the trial of a new-economy case a daunting challenge to the factfinding capacity of the judiciary. Nor are the economic issues that I have identified issues that can be left to a neutral expert to resolve, given the many imponderables they involve; they are not problems simply to be solved by the application of the relevant economic principles.

The rapidity of innovation in the new economy has another very important institutional implication. Federal courts are highly efficient by the standards of the American legal system. The federal court queue is short, and strong district judges can move even complex cases along briskly. But this is speaking relatively. Antitrust litigation moves very slowly relative to the new economy. Law time is not real time. The law is committed to principles of due process that limit the scope for summary proceedings, and the fact that litigation is conducted by lawyers before tribunals that are not technically trained or experienced inevitably slows the process.

The mismatch between law time and new-economy real time is troubling in two respects. First, an antitrust case involving a new-

economy firm may drag on for so long relative to the changing conditions of the industry as to become irrelevant, ineffectual. That was a problem even in the old economy. One recalls for example that by the time the monopolization case against Alcoa completed its journey through the courts, Alcoa had lost its monopoly for reasons unrelated to the litigation; as a result, the decree finally entered against Alcoa offered little more than nominal relief (the divestiture of Alcoa's Canadian subsidiary, a substantial producer but one which U.S. tariffs prevented from having a big impact on the U.S. market[17]). This type of problem is likely to be more frequent in the new economy.

Second, even if the case is not obsoleted by passage of time, its pendency may cast a pall over parties to and affected by the litigation, making investment riskier and complicating business planning.

These problems are aggravated by the tendency of antitrust litigation to create multiple lawsuits out of a single dispute; call this the cluster-bomb effect. No sooner does the Antitrust Division bring a case, but the states and now the European Union are likely to join the fray, followed at a distance by the antitrust plaintiffs' class-action bar. The effect is to lengthen out the original lawsuit, complicate settlement, magnify and protract the uncertainty engendered by the litigation, and increase litigation costs.

I wish I had a good solution to these problems, or at least some of them, but I don't. As minor palliatives, I would like to see greater use of neutral experts, as I have explained, and I would also like the Antitrust Division and the Federal Trade Commission to be given the necessary appropriations to enable each of these agencies to hire a competent technical staff.[18] That won't be easy, given the salaries that competent new-economy scientists and engineers command in the private market. But there is a more serious problem, and that is the rapid obsolescence of technical knowledge, which combined with tenure practices in government agencies may prevent the agencies from maintaining a technical staff that is actually abreast of current technology. This is an objection, by the way, to an alternative that I consider distinctly unpromising, which is to have a specialized, technical court composed of judges with technical training. Their training

---

[17] See See Merton J. Peck, *Competition in the Aluminum Industry 1945–1958* (1961), under index references to Limited. (The name of the Canadian subsidiary was "Aluminium Limited.")

[18] I am not proposing a net increase in the appropriations for these agencies, just that they be authorized to allocate a portion of their existing appropriations to hiring technical experts.

is likely to become out of date before they become judges, unless we are to have a system in which the regulation of the new economy is entrusted to 25 year old judges, lawyers, and engineers.

Now turning to more radical measures, I would like to see, first, the states stripped of their authority to bring antitrust suits, federal or state,[19] except under circumstances in which a private firm would be able to sue, as where the state is suing firms that are fixing the prices of goods or services that they sell to the state. (In other words, only their power to bring *parens patriae* suits would be abrogated.) States do not have the resources to do more than free ride on federal antitrust litigation, complicating its resolution; in addition they are too subject to influence by interest groups that may represent a potential antitrust defendant's competitors. This is a particular concern when the defendant is located in one state and one of its competitors in another, and the competitor, who is pressing his state's attorney general to bring suit, is a major political force in that state. A situation in which the benefits of government action are concentrated in one state and the costs in other states is a recipe for irresponsible state action. This is a genuine downside of federalism. The federal government, having a larger and more diverse constituency, is, as James Madison recognized in arguing for the benefits of a large republic, less subject to takeover by a faction. I am not myself inclined to make a fetish of federalism.

When I was a law clerk at the Supreme Court, almost 40 years ago, I was struck by the poor quality of the briefs and arguments of the lawyers in the offices of the state attorneys general. Since becoming a judge almost 20 years ago, I have been struck by the poor quality of the briefs and arguments of most though not all of the lawyers in the offices of the state attorneys general of my circuit. On several occasions our court has found it necessary in published opinions to criticize the performance of these lawyers. In response to one of these opinions a meeting was arranged between the Attorney General of Illinois and myself (I was chief judge then), at which he explained that the very low salary scale of his office makes it difficult to hire competent lawyers.

One of the sources of the quality problem is that in most states the attorney general runs for office independent of the governor and is thus not part of the governor's administration and as a result does

---

[19] That is, Congress should enact a statute that would make antitrust law, like patent, copyright, and bankruptcy law, an exclusively federal body of law enforceable only in federal courts.

not receive much support for the governor in negotiating for appropriations from the legislature. In addition, in a field like antitrust, the offices of the state attorneys general are not able to scale up to anything like the level of specialization and professionalism of the Antitrust Division of the U.S. Department of Justice, with its hundreds of lawyers and economists.[20]

A possible, partial solution to the delay and cluster-bomb problems would be to empower the Justice Department to bring antitrust suits for damages by whomever suffered and to parcel out the proceeds of the suit among the victims. If the Department decided to bring a suit, private suits (and state suits, if they were still permitted) would be preempted. In effect the Department would have a right of first refusal, which it presumably would exercise in major cases in order to consolidate and thus simplify and expedite the litigation. This is not an original suggestion, as it is a power already enjoyed by government agencies such as the Department of Labor and the Equal Employment Opportunities Commission in the field of employment law.[21]

Any suggestion for streamlining antitrust enforcement is likely to be met, from the conservative side of the policy-analysis spectrum, by the argument that it is a mistake to make government more efficient, when efficiency connotes the better matching of means to ends and the ends may be bad ones. I raised this possibility in discussing antitrust enforcement by the states. I must confront it with regard to the federal agencies as well. The essential question is the degree to which these agencies are buffeted by interest-group pressures, as I believe the state antitrust enforcers are. I think that, today, only to a small degree. There is a history of efforts to explain antitrust enforcement as just another example of interest-group politics, an approach that has worked for a number of government agencies and programs. But the efforts have not been successful,[22] and the reasons

---

[20] Cf. Robert H. Lande, "When Should States Challenge Mergers: A Proposed Federal/State Balance," 35 *New York Law School Law Review* 1047, 1064–1066 (1990).

[21] See, for example, 29 U.S.C. § 626(c)(1); EEOC v. G-K-G, Inc., 39 F.3d 740, 744–745 (7th Cir. 1994).

[22] See, for example, William F. Baxter, *The Political Economy of Antitrust* (Robert D. Tollison ed. 1980); George J. Stigler, "The Origin of the Sherman Act," 14 *Journal of Legal Studies* 1 (1985); Malcolm B. Coate, Richard S. Higgins, and Fred S. McChesney, "Bureaucracy and Politics in FTC Merger Challenges," in *The Causes and Consequences of Antitrust: The Public-*

may be that the agencies are dominated by lawyers most of whom go on to service in the private sector and that antitrust law itself is dominated by federal judges exercising a broad discretion because of the open-endedness of the major federal antitrust statutes. To land good berths in the private practice of law the antitrust enforcers must demonstrate their professionalism, which means keeping within the boundaries fixed by the courts. Federal judges with their secure tenure seem largely insulated from the interest-group pressures that play on the other branches of government. Looking over the entire history of U.S. antitrust law, I conclude that the most powerful explanatory variable is simply the state of economic opinion. Antitrust doctrine has changed more or less in tandem with changes in economic theory, albeit with a lag.[23] As a result, I am not fearful of making antitrust enforcement more efficient.

Unfortunately, the measures that I have suggested, even in the unlikely event that they are adopted, would probably not do a great deal to correct what seems to me a serious mismatch between the conditions of the new economy and the institutional structure of antitrust enforcement. That brings me to the final question I address, which is society's proper response to a situation of at least temporarily ineradicable uncertainty concerning the effect of governmental intervention in the economy. We really don't know what the effect of applying antitrust principles to the new economy will be, except when they are applied just to stop horizontal price-fixing or mergers of major competitors in highly concentrated markets. I think a policy of zero enforcement against alleged exclusionary practices in the new economy would be a mistake, because there is a pretty solid theoretical basis for concern both that some new-economy firms would find it in their rational self-interest to employ such practices and that natu-

---

Choice Perspective 213 (Fred S. McChesney and William F. Shughart II eds. 1995). But see Roger L. Faith, Donald R. Leavens, and Robert D. Tollison, "Antitrust Pork Barrel," 25 *Journal of Law and Economics* 329 (1982), finding evidence that the FTC in the 1960s was unduly influenced in its enforcement activity by the parochial interests of members of Congress who had budgetary and oversight powers with respect to the Commission.

[23] As Louis Kaplow has pointed out, it is not true that the federal courts in antitrust cases suddenly discovered economics in the 1970s. Kaplow, "Antitrust, Law and Economics, and the Courts," *Law and Contemporary Problems*, Autumn 1987, pp. 181, 187. See also Herbert Hovenkamp, "The Antitrust Movement and the Rise of Industrial Organization " 68 *Texas Law Review* 105 (1989).

ral market forces would not undo those practices in time to avoid significant social costs. A policy of zero enforcement would also deprive us of important information about competition and monopoly in this vital sector of the national economy. Cleary, though, the byword of a prudent enforcement agency and a sensible court will be: caution.