Chapter 3

CONCEPTS OF COMPETITION AND EFFICIENT POLICY IN THE TELECOMMUNICATIONS SECTOR

WILLIAM G. SHEPHERD*

Two sets of events related to the point-to-point telecommunications sector have recently bred expectations of radical new ideas. First, AT&T's settlement with the Antitrust Division of the U.S. Justice Department offers large practical changes in its organization. Second, several Bell-related economists have offered new ideas ("sustainability" and "contestability" of monopoly) which are claimed to replace the mainstream literature on competition and to overturn the accepted lessons for public policies.

On both levels there is currently much flux. In this article, I can only aspire to address the deeper issues rather than the latest developments. My argument will be critical, both of the antitrust settlement and of the new theories. But my main conclusion is also highly positive: the basic theory of competition is both sound and effective in guiding the main revisions that the telecommunications sector may need.

---

* William G. Shepherd is Professor of Economics at the University of Michigan, where he has been investigating industrial organization and regulatory issues. Formerly he was the Special Economic Assistant to the Assistant Attorney General for Antitrust at the U.S. Department of Justice, and was a Research Scholar at the Brookings Institution. He has been a frequent advisor and witness in antitrust and regulatory cases. He is the author of several books and many journal articles. He received a BA from Amherst College and an MA and PhD from Yale.

In the first section, I will cover a variety of recent developments in theory, policy, and real-world conditions. Next I will consider the economic choices posed by a Bell-like system, with its several vertical and geographical parts, varying degrees of viable competition, and problems of access by competitors. The third section will consider the complex problems in moving from monopoly to competition. Then I will assess "sustainability" and "contestability." The following section will address the organization of the sector itself, offering several suggestions for revising the antitrust settlement and treating future problems.

BACKGROUND

A series of recent research findings are relevant in judging the telecommunications sector.

The Increase in Competition in The U.S. Economy

Since 1960, there has been a marked increase in competition throughout most of the U.S. economy.[1] The share of pure monopolies and dominant firms has dropped from 8 percent of national income to 5 percent, while the share of tight oligopoly has declined by half, from 36 percent to 18 percent. The increases in competition have occurred in every major sector. Altogether, the total share of markets with effective competition has risen from 56 to 76 percent of national income.

In this new setting, cases of high market power—such as in telecommunications—stand out as distinct deviations which invite corrective treatment.

The Crucial Role of Antitrust

A large share of the rise in competition has been caused by antitrust activities and deregulation, which itself has largely been a response to antitrust pressures.[2] Though often debatable in specific cases, antitrust activities encompass a wide range of informal and formal actions which have had a large total effect. The rise in competition indicates the value of maintaining these actions toward the

remaining instances of market power. Despite recent denials by Chicago-oriented economists, the economic, legal, and social support for strict antitrust enforcement is strong.

The Receding Importance of Economies of Scale

Technical scale economies have not been a major cause of market power in the last two decades, and they appear generally to be receding rather than rising in importance.[3] That large trend reflects several forces, including the growing use of computer technology, electric power equipment, and changing worker attitudes. In point-to-point telecommunications, too, the economies of scale (and economies of scope) now appear to be dwindling compared to market size in many parts of the sector.

Improvements in Pricing Criteria and Practice

The basis for utility pricing has become markedly more refined in fitting marginal costs, both in electricity and telecommunications. The ability to define and achieve efficient price structures, and to identify predatory or other unfair pricing strategies, has increased.

Changes in Telecommunications Technology and Demand

Large changes are occurring in many parts of the telecommunications sector, including switching, transmission, and the composition of traffic (between voice and data). Moreover, there is an increasing overlay with adjacent markets, especially cable television, data processing, and electronic equipment. Therefore some parts of the sector have become "naturally competitive," with minimum efficient scale small enough to permit at least several viable competitors of similar size. Many other parts are expected to become naturally competitive (or at least partially so) in coming years.

Taken together, these changes call for replacing regulated monopoly with effective competition throughout most of the sector. But there are crucial questions of degree and timing. Deregulation is a

process which should occur *as* competition becomes effective, not before. But how much competition is enough to be effective, and how can it be identified and measured? Because the Bell system is the long-established monopolist in most of the sector, that system's officials naturally prefer to have deregulation immediately, and Bell-associated theorists have prepared a rationale for withdrawing regulation quickly.

DEFINING COMPETITION AND STANDARDS OF PERFORMANCE

Before considering that rationale, I will first briefly set out the basic concepts of competition and efficiency since there has recently arisen some confusion over these matters.[4] The following summary is meant to reflect the established core of research in the field.

What conditions determine the degree of competition? One begins with the structure of the market, but not because it precisely determines behavior and performance. It doesn't. But structure does usually have strong effects on market outcomes, and its main elements can be measured fairly reliably, more reliably than the behavioral aspects.

Market Structure and Actual Competition

The primary element of structure is the leading firms' market shares. *Ceteris paribus,* the larger the share of the largest firm, the higher the degree of market power, as illustrated in Figure 1. The extra profits can be attributed either to the simple raising of price, or to price discrimination, or to both. (Economies of scale are usually a minor cause of the extra profits (Shepherd, 1982).) Firms with high market shares usually have high market power and opportunities for excess profits; firms with low market shares do not, even if they are in the same market.

A small market share gives the firm little market power or influence. That applies also to new entrants to a market; small entry usually has little or no competitive effect. The degree of effect may vary with the structure of the existing firms; for example, a 2-percent entrant might have more effect on a prior pure monopolist than on ten firms with 10-percent market shares. In fact, the whole prob-

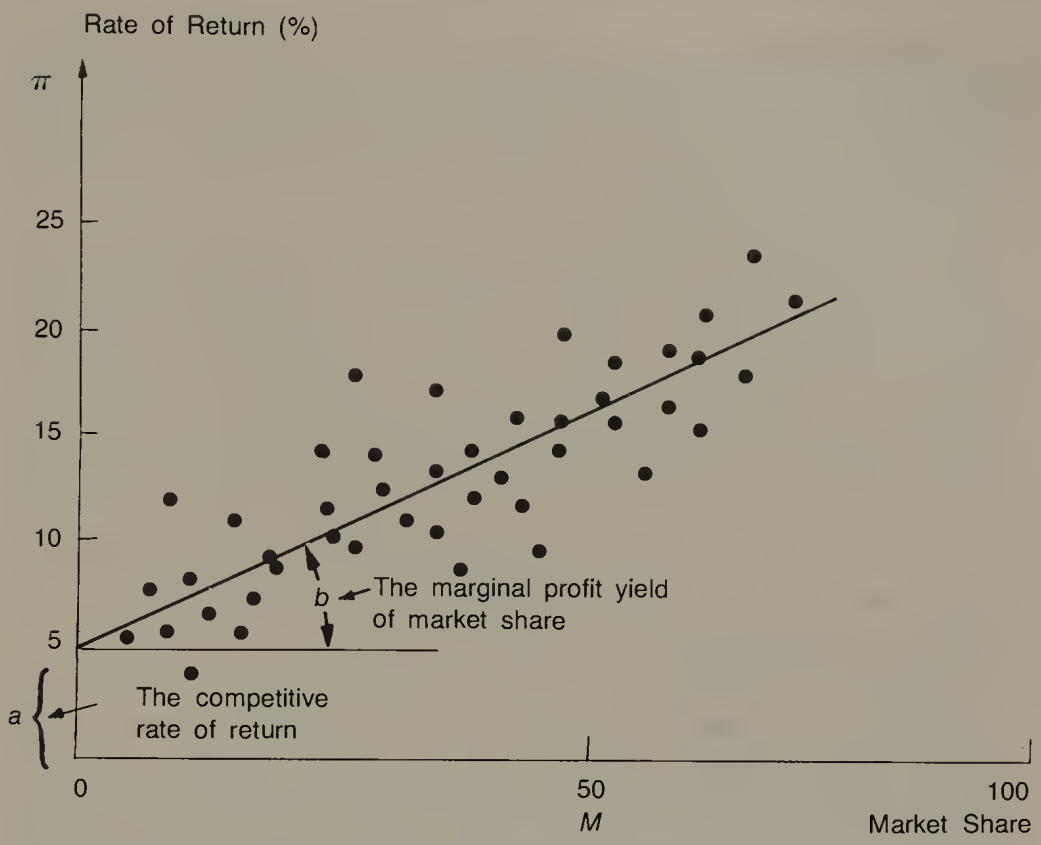

Figure 1. The basic relationship between market share and profitability (illustrated here in linear form). Copyright © 1979 by permission of Prentice-Hall, Inc., Englewood Cliffs, N.J.

lem of measuring the degree of entry and its effects has been largely undeveloped ever since Joe S. Bain's (1956) initial work.

Nonetheless, it is well established that firms with small market shares (including entrants) have small effects on the market. Conversely, firms with dominant market shares (above 40 to 50 percent, with no close rival) have high market power, except in unusual situations. More generally: *market dominance is inconsistent with effective competition*. An implication here is that a dominant firm's market share will need to be lowered below 50 percent and possibly lower before effective competition can be trusted to apply constraints in formerly regulated markets.

Actual competition in the market—as shown by the market shares and other conditions of established firms—is crucial. It exists and exerts force. In contrast, *potential competition (the possibility that some firm not now in the market will enter at some future time) is usually a secondary and often a trivial matter compared to actual competition*. Entry is a probabilistic thing, of unknown degree, both

in its likelihood and in its size. For example, a potential entrant that is expected (with a 50-percent probability of the entry occurring and succeeding) to try to achieve 8 percent of the market is equivalent to a firm that is already in the market holding a 4-percent market share and trying to double it.

On the whole, the recent preoccupation of some analysts with entry is a strange effort to replace the heart of the problem with a literally peripheral aspect. This deviation is especially unfortunate because the measurement of both entry and entry barriers has remained in a primitive state, far less accurate than the measurement of actual structure within the market. Accordingly, there has been a high ratio of optimistic theorizing about entry compared to solid evidence showing what it is and that it actually has strong effects. Virtually no evidence exists to show that potential entry is more important than actual competition; the weight of research is overwhelmingly the other way.

One would not need to stress that potential entry's role is peripheral and unproven, had it not recently evoked so many increasingly grandiose claims. To be sure, entry is not always of slight concern; some specific barriers (such as patents) can be crucial in individual cases. But generally, entry is correctly seen as only a possible modifier of the more central conditions that exist within the market.

Several other basic points about competition need to be noted. One is that competition tends to raise X-efficiency and the rate of innovation, often by amounts that dwarf its welfare effects on allocative efficiency. Therefore, static models focused on the maximizing of consumers' surplus under simple assumptions are always incomplete, and often they are deceptive in purporting to evaluate competition and its total social effects.

Another point is that the market shares of leading firms in actual markets are normally well above (often by a multiple in the case of dominant firms) the size required by minimum efficient scale (MES).[5] This *surplus market share* is socially costly, because it provides unnecessary market power.

The Learning Process

In considering telecommunications, I need also to stress that the process of learning, as embodied in learning curves, can affect competition. A learning-curve situation gives a cost advantage to the

firm that is first (and furthest) to move down the curve. This firm may be able to slow down or prevent its competitors' move down the curve to catch up with the leader. Any advantage this firm may have in controlling technology and/or the learning process will reduce competition. Moreover, this advantage is increased if the underlying technology is advancing. Then the leading firm's ability to extract extra profits and limit its competitors' ability to catch up will be larger. Once the firm is ahead, it may be able to stay ahead, and not because of any innate superiority. On the contrary, that firm's leadership may reduce the rate and diffusion of technological progress below the results that effective competition would yield.

Marginal Cost

Finally, the efficient basis for prices (and price structures) is marginal costs. In a well-planned firm, this will not cause a problem of financial deficits. Therefore, price discrimination is not justified socially in the usual cases. Only in extreme cases of excess capacity or overhead costs will a set of marginal-cost-based prices cause a deficit. Even in such unusual cases, the price discrimination may not need to depart more than slightly from marginal-cost pricing patterns. Conversely, because discriminatory prices do not usually offer social gains, competitive pressures away from discrimination and toward cost-based prices need not be socially harmful.

CONCEPTS FOR THE TRANSITION FROM UTILITY TO COMPETITION

With these points as a basis, I will now analyze the shift from regulated monopoly to deregulated competition. That transition is more complicated than the literature has recognized, and it presents several specific problems that I will address. With this review in hand, we can then consider the Bell economists' ideas and the actual treatment of telecommunications.

The concepts to be considered here are the "room" for competition, core services, separability, joint costs, and responsibility for service.

Regulation is a distinct contractual situation, giving the firm certain rights, duties, and protections. Introducing competition can change the firm's role, usually in complex ways. The optimal treat-

ment in traveling from regulated monopoly to partial or complete competition depends on certain economic conditions (of costs, separability, and competition) which are discussed in this section. It is necessary to begin with basic issues of regulatory economics, for that is what the sustainability literature has set out to recast.

Conventional Natural Monopoly and Ideal Regulation

The conventional regulated monopoly is defined by conditions of demand and cost.[6] If market demand is small relative to economies of scale, as in Figure 2, one observes the simple case of natural monopoly. Demand may intersect with average cost in the declining part (as does D_1) or just where average cost reaches its minimum, as does D_2. The latter part is called minimum efficient scale (MES). In either case, D_1 or D_2, competition could not survive. The monopolist, with its marginal revenue (MR), will prefer to set one output at Q_M (where MR_1 equals Marginal Cost), and one monopoly price such as P_M, earning monopoly profits as shown. It may also try to segment the market and set discriminatory prices, in line with customers' varying demand elasticities.[7] That can yield still higher profits.

Under regulation, the monopoly is usually given a public franchise as the exclusive supplier, and regulation is then supposed to hold price down to average cost—at P_R (*R* for regulation) for demand of D_1. Ideally, price would be set equal to marginal cost. But with a demand of D_1, marginal-cost pricing would cause financial deficits, as shown. That long-familiar analytical problem requires a "second-best" solution. Some analysts still consider it to be *the* utility problem, inexorably imposing difficult choices.

But that often need not be so. Demand may grow to D_2 or beyond, or costs may evolve toward smaller scale economies; or both may occur. A properly planned growth path for the utility will keep the firm's capacity in line with demand, so that demand intersects the average cost curve at its minimum, as does D_2 in Figure 2.[8] Then price can equal marginal cost while yielding zero or positive profits, and the problem of the deficit disappears. Moreover, cost variations in multiproduct firms (including demand fluctuations between peak and off-peak times) can call for a structure of prices

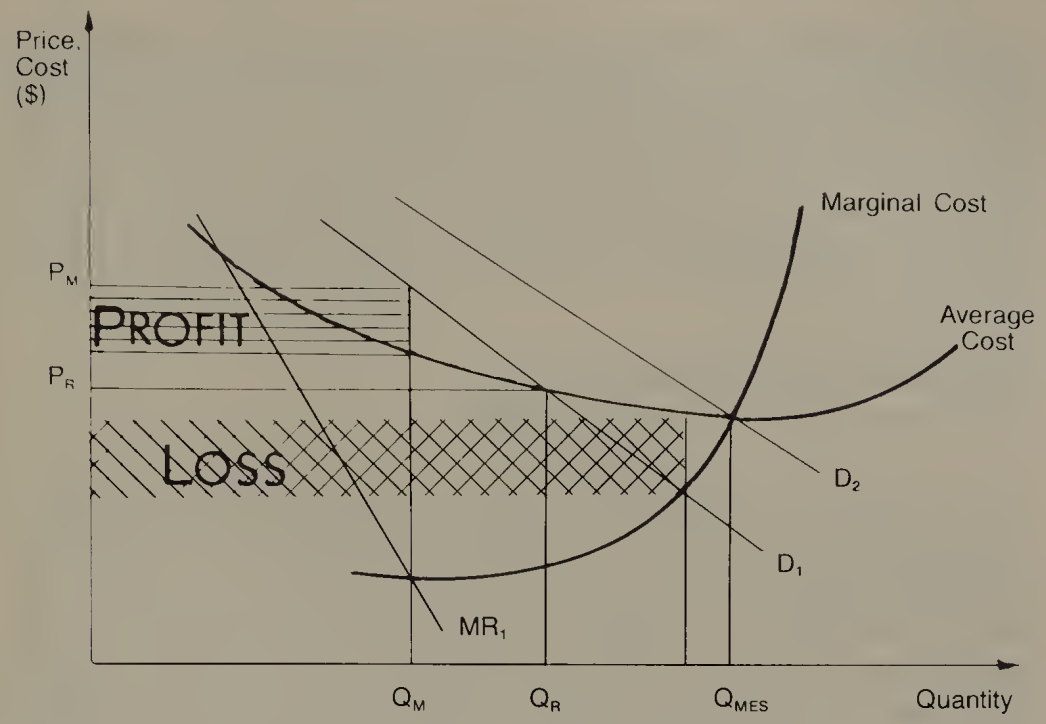

Figure 2. Simple natural monopoly under regulation.

in line with the marginal costs. Under common conditions, this array of prices can still yield sufficient profits. Regulation prevents any unnecessary price discrimination.

The franchised firm bears the responsibility for service in two senses: (a) it must extend service to all who meet the going price. That is the conventional "common carrier" duty; (b) the firm is also supposed to guarantee all customers reliable service: its capacity will be adequate and assured. Because customers have to rely on this supplier, their activities come to depend on its continuous, ample supply. A service failure therefore imposes severe costs and imperils the firm's retention of its franchise.

In the ideal static case (with D_2 in Figure 2), perfect regulation forces price down to average cost, permitting just adequate profits. No large excess profits are permitted, and so major windfall capital gains are prevented. But the regulators accept a duty to protect the utility's asset values from harm. Investments made in good faith and approved by the regulators will be protected by the regulators' decisions on "fair" prices and profits. Management is assumed to be both internally efficient ("X-efficient") and quick to innovate all worthwhile new techniques and products.

Such perfection never occurs in this world, and actual cases often depart sharply from it. If regulation is ineffective, then the firm may move to classic monopoly behavior: especially (a) excess profits, (b) excessive price discrimination, (c) X-inefficiency, and/or (d) retarded innovation. Various side effects of regulation may also occur, including an enlargement of the firm's investment (the Averch-Johnson rate-base effect).

Room for Competition and the Cross-over Zone of Market Share

Now consider the advent of possible competition. Demand may shift out beyond the given declining part of the average cost curve; or new technology may reduce minimum efficient scale; or both changes may occur in parallel. Where demand intersects the flat average cost curve, as in Figure 3, there can occur the "first-best" output level, with price equal to marginal cost and no excess profits. Call that level Q_I (*I* for industry output). Between MES and Q_I, there is now room for competition. That "room" may be of any degree, from small to large, depending on costs and demand.

Figure 3 illustrates two interesting cases. First, with demand D_3 there is some slight room for competition (MES to Q_{I_3}), but the original monopolist's market share at MES will continue at over 50 percent. Such cases—where M_O remains between 50 and 100 percent of the market—are natural quasi-monopoly (or the "dominant firm" case). The new "room" for competition (between MES and Q_{I_3}) may, of course, be captured by the established monopolist, rather than by new entrants. The actual market share taken by new entrants is indeterminate: a number of other conditions may influence it. Though natural quasi-monopoly conditions make some competition possible, they do not assure that competition will actually emerge, nor that it will be effective.

Second, further relative demand growth may give natural oligopoly, as illustrated by demand curve D_4. There is now room for four equal-size firms, and competition could become severe. Yet this tight oligopoly might instead develop collusion, either tacit or direct. And the original monopoly may, as before, be able to capture all or part of the output above MES for itself, with M_O staying as high as 100 percent.

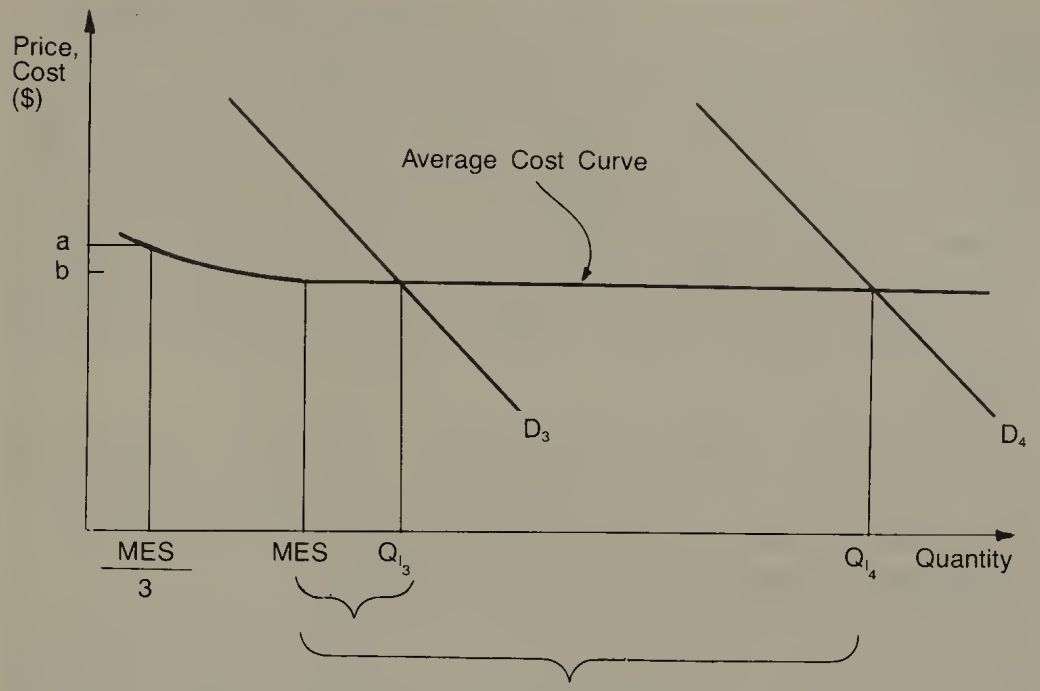

Figure 3. Varying "Room for Competition."

Finally, if the demand and cost shifts go still further, then natural competition may emerge, with an even larger ratio of Q_I to MES. Many efficient firms might be able to coexist, and price will be driven down to average cost. The original monopolist may find it hard to retain a high market share; yet conceivably its M_O could still remain high.

The spectrum therefore contains all degrees, ranging from pure monopoly, quasi-monopoly, and oligopoly to full competition. Evidently, the shift from traditional natural monopoly regulation to natural competition—which permits full deregulation—is a large one. It presumes a radical relative shift of demand and of scale economies. Any lesser shift, say to Q_I at double the level of MES, will leave a high degree of market power even if M_O is held strictly to MES. Deregulation will need to be a careful, complex sequence—dealing with pricing, entry, mergers, and other aspects—so as to fit and use the growing room for competition. No easy, quick deregulation is appropriate. Moreover, regulation will often lack complete information on these conditions, and will have to guess at them.

The shift toward competition involves a cross-over zone, defined in terms of the market share of the established firm, M_O. As M_O

declines across the range, effective regulatory constraint may be replaced by the constraints of a competitive market.

The cross-over zone of M_0 is not a unique level of market share. It depends on other firms' market shares and on the height of barriers to further entry. The literature on market structure and performance suggests that the zone usually lies between about 60 and 30 percent of the market.[9] With an M_0 of 60 percent or higher, competition cannot be fully effective; but when M_0 is below 30 percent and entry barriers are low, competition usually is fully effective.

The benefits of effective competition are the same as those for ideal regulation; prices aligned with costs, X-efficiency, and innovation. If M_0 remains above the cross-over zone, then competition is inadequate and regulation needs to be retained.

Related to the cross-over zone is the minimum efficient market share. It is the M_0 value that provides an optimum balance between (a) the gradient of the average cost curve below MES and (b) the efficiency and innovation benefits of competition (Scherer (1980), Shepherd (1982)). The cost gradient is commonly defined as the cost penalty from being only one third the size of minimum efficient scale, as a percent of minimum cost. That point is illustrated in Figure 3 at MES/3, where average cost of *a* is about 15 percent above the lowest attainable cost *b*. The cost gradient is:

$$G = \frac{a-b}{b}$$

MES does not set a rigid lower border on M_0. If the cost gradient is low, then setting M_0 below MES will raise cost only slightly, while the added competition may yield more X-efficiency and innovation. The relationship between M_0, MES, and G is complex, and deserving of further research; on its basic form, see my paper in Sichel and Gies (1982).

Core Services

Most utility sectors contain certain operations that are regarded as, or merely claimed to be, the essential core or network that makes the whole system's service possible. Other services are adjacent: ancillary, normal, or ordinary, in some sense.

Core services are not uniquely defined. They may be those which have the greatest relative economies of scale. Often they share joint costs of production. But scale economies and joint costs are neither necessary nor sufficient. Core services may simply be regarded as the unifying or crucial part of the system. For example, local telephone switching service may be a natural monopoly. But telephone officials may instead regard intercity switching and transmission capacity as the real core of their system, essential to its unified operations. Several other instances of core services are given in Table 1. Evidently, core services are often inputs to the utility system. To that extent they occupy a separate vertical stage of production and can be separated from the utility firm.

Core operations are often controversial, for the established forms have an interest in contending that the core is large, even if it is not. Core services are usually the main part in any argument that the existing monopolist needs to be protected. If there are no core outputs, then it does not matter who supplies which outputs. And if consumers value core services at less than their costs, then such services can be omitted altogether.

I am not urging the importance nor even the existence of core outputs; on the contrary. But since the recent "sustainability" analysis does need them, it is important to consider them explicitly. Though as this stage they have not been fully defined, we can analyze their general role.

Separability

The crucial question is how separable the utility's services are. If they are closely mingled—with joint production and high overhead costs, often called "economies of scope"—then their provision will need to be done by a single unified multiproduct monopolist. But if instead production and costs can be cleanly separated for each output—with all costs fully allocable because there are no economies of joint production—then each output can be treated as in a separate market. Products with natural monopoly conditions can be separated from the rest, and competition can be applied just where it is appropriate.

The degree of separability is often unclear, particularly to regulators and other outside observers. Technology is usually not transpar-

Table 1. Examples of Possible "Core" Services.

Sector	Possible "Core" Services	Presently Provided by
Airlines, air freight	Airports	Public agencies
	Flight guidance systems	Federal Aviation Authority
	Reservations systems	Each Airline
Trucking, intercity buses	Highways and roads	Public agencies
Satellite communications	Satellite launching systems	National Aeronautics and Space Administration
Telephones	Local switching and loops	Telephone companies
	Intercity routing (?)	Bell System
Postal	Local delivery routes	Postal System
	Local sorting	Postal System
Electric power	Generation (?)	Electric firms
	Local distribution	Electric firms (private and public)
	Regional pooling	By agreement among electric systems
Banking	Check clearing	Federal Reserve System

ent, especially for complex monopoly systems such as telephones. Moreover, separability is often not just a technical given. The firm may be able to control the degree of separability among its outputs, in two different ways. First, it can select the technologies that have the largest element of joint production and overhead cost, avoiding the more highly separable ones. Second, it can conduct its operations and cost accounting in ways that make the joint costs and overhead costs seem as large as possible.

When that occurs, the recorded marginal costs for individual outputs will be minimized, giving the utility the widest discretion in its pricing actions, and in its pricing reactions to new competitors. The utility has an interest in minimizing the apparent degree of separability, so as to make its monopoly control of most or all services seem necessary. Often the firm has wide discretion in doing so. The result can be to overstate sharply the true degree of "economies of scope" within the utility "system."

Where true economies of scope *are* large, there may exist a set of true core services. (Yet even so, they may be a peripheral subset of outputs, interrelated but a minor part of the whole set of outputs.) Some core services can be provided by the utility or separately by some other unit, perhaps a public agency (such as the public provision of airports and FAS flight guidance systems).

At any rate, core services (if they do exist as such) can often be quite separate in provision and pricing from other outputs of the sector. If public policy succeeds in maximizing that separability, then the viability of competition in quasi-monopoly settings will be enhanced. Table 1 offers some examples showing that separability can be complete. "Core" services in several sectors (airlines, air freight, trucking, and banking) are provided separately from the firms themselves.

Since core services do not necessarily have large scale economies, they need not create a natural monopoly. Competition may be possible, and firms may be able to proliferate, each with its own core services. Airline reservations systems are an example of that; each airline is able to provide itself with the core system.

Core conditions may change or fade away, as technology develops. There have been marked reductions of core conditions in a variety of sectors. Indeed, the recent deregulation series in various sectors has mainly occurred following the shrinkage or disap-

pearance of core conditions. Therefore, the regulated multiproduct monopoly bears a burden of proof to demonstrate that its claimed core conditions and complementarities of production do exist, and that they are unavoidable by alternative technological choices. Otherwise the regulators, with their limited knowledge of technology and internal company conditions, may (a) be unable to find the true conditions, (b) be led to underestimate separability, and therefore (c) be maneuvered so as to limit competition unnecessarily.

Joint Costs and Entry

These concepts help to redefine the role of joint costs in the face of entry by low-cost firms. Consider the simplest case, with two outputs 1 and 2 produced jointly by monopolist A. There are separable marginal costs for each output, plus some shared overhead cost.

Now let a new firm appear, able to produce output 1 at *less* than its marginal cost to the monopolist. *At that point, the output ceases to possess the economic attribute of a joint product with output 2.* As a joint product, it can no longer survive. Therefore output 1 will henceforth be produced only with the new method, by a new firm B, and/or by the existing monopoly firm A. The former joint costs must now be assigned by firm A solely to output 2. Whether to produce output 2 depends on the demand conditions. If consumers are not willing to pay the combined cost, then either: (a) the output is to be discontinued, or (b) the overhead costs are (partly or wholly) written off, reflecting the new true cost of output 2. Firm A suffers a loss of asset value, but this is merely the efficient outcome of technological progress in producing output 1.

In this way, innovation in producing individual outputs may reduce the breadth of the monopolist's offerings, unless the monopolist preempts entry by adopting the new technology first. The innovation may force up the monopolist's apparent costs on the remaining outputs. Yet the shift of overhead cost to output 2 is part of an efficient assimilation of new technology. Only by letting it happen under the threat or reality of entry can the innovation (by either firm) be assured. The monopolist's financial injury is inevitable. In practice, of course, the injury can be lightened or avoided, if the monopolist is far-sighted and progressive, so that it avoids mistaken commitments to doomed technology.

The Responsibility for Service

Under conventional regulation, the franchised firm has borne exclusive responsibility for service. Any shrinkage of that domain of responsibility seems to challenge the firm's exclusive role, as well as the interests of customers in having reliable service. New competition is therefore often resisted for being likely to undermine reliable service. *Yet eliminating or diluting that exclusive responsibility for service is necessary if competition is to become effective*. In virtually all industries (regulated and unregulated), goods and services are provided steadily and reliably by several or many competitors, even though no one firm is required to accept any formal responsibility for service. Competition and profit seeking induce an adequate supply to be forthcoming. Of course there may be occasional lapses in supply. But such lapses also occur occasionally in exclusively "responsible" utilities.[10] The idea of absolute responsibility of franchised utilities for service is largely a myth. Accordingly, competition's dilution of responsibility is merely a matter of degree, and *actual* service reliability may increase. Only in extreme cases of certain emergency services is absolute responsibility possible and perhaps desirable.

Furthermore, customers can usually adjust so as to be less reliant on utility service. Industrial customers often have their own reserve power sources, to tide over electricity failures. Families can often adjust their consumption habits and appliances so as to be less vulnerable. Such adjustments to lighten the future impacts may impose higher costs, but the net rise in costs may be small.

In short, we already live in a world of *partial* service responsibility. Competition may increase reliability, not reduce it. Even if reliability is reduced, the costs are finite and may be small.

THE ANALYSIS OF SUSTAINABILITY

With these points in mind, we can now consider the question of "sustainability." Since 1976, Baumol, Bailey, Willig, Panzar, (BBWP) and others have written a series of papers on the "sustainability" of regulated utility firms facing new competition.[11] Writing in highly abstract terms, they have suggested (a) that regulated firms tend to adopt optimal prices and outputs, and (b) that new

competition will often make those prices and outputs (and possibly the entire firm) "unsustainable." Though the authors occasionally disclaim any policy relevance of their analysis, they and many others frequently describe it as important and authoritative. So it is timely to assess "sustainability," even if only briefly here.[12]

I will point out that the issues now labeled "sustainability" have been long debated, using different terms and methods. The crucial functions and outcomes are continuous, not of a disjunct, checkerboard nature as BBWP suggest. The competition that may occur is a matter of degree, across a familiar spectrum of conditions, which may or may not entail significant welfare losses. In this light, the sustainability literature has yet to offer important lessons about the familiar basic issues.

BBWP present the regulated monopoly firm as being fragile and vulnerable to entry. The firm is likely to reach an optimal result, they indicate, but only if it is protected from entry. In taking this line, BBWP ignore the enormous literature which establishes that utilities are strong and often wasteful under regulation and that competition can often improve their economic performance.[13]

A Summary of BBWP Analysis

The main assumptions of BBWP analysis are given in the left-hand side of Table 2. Most BBWP articles have offered static analysis, but BBWP also offer a two-period "dynamic" discussion, based on construction costs in the second period. The assumptions for that are included in the second part of Table 2.

Regulation is assumed by BBWP to be "ideal": omniscient and all-powerful. Optimal prices and outputs are known and enforced. This result holds even if the regulated firm is a multiproduct producer with complex joint-production and overhead costs.

Under natural monopoly conditions (assumption 2) and ideal regulation, the firm offers the ideal set of products, at "Ramsey" prices which maximize consumer surplus.[14] Since there is unlimited reselling of products (assumption 3), the monopolist can set only one price for each product. BBWP do not consider the nature or relative importance of those various products (core or adjacent; large or small volume). Nor do BBWP permit the firm or new entrants to vary the products as the action proceeds. By assumptions 5 through

Table 2. Assumptions for BBWP Sustainability Analysis.

BBWP Assumptions	Common Conditions in Reality
I. Static Analysis	
1. *Regulation is ideal.* Regulators' information and powers are complete, and their actions impose ideal results without any friction or delay. The monopolist produces the socially ideal set of products, under regulation. The monopoly earns zero excess profits.	1. Actual regulation is usually imperfect, incomplete, and lagged. The firm has wide discretion in product offerings. Profits are often well above zero.
2. *Costs.* For the single-product firm, there are large economies of scale. For multiple-product firms, there are high joint and overhead costs.	2. Cost conditions vary, in both scale economies and joint costs.
3. *Reselling.* All products can be freely resold among customers, so that only one price can be set by the monopolist for each product.	3. Reselling is usually not feasible, except for certain limited products. Price discrimination is usually possible.
4. *Nature of the products.* (a) No distinction is made between core and adjacent goods. All are treated as if	4. The products vary, both among core and adjacent services and between the established firm's products and

Table 2. *(Continued.)*

BBWP Assumptions	Common Conditions in Reality
I. Static Analysis	
they were core products. (b) The monopolist's and entrants' goods are identical items; no variation is permitted, either at the start or as events unfold. (c) The social value of goods is reflected only in demand conditions (including consumer surplus).	entrants' products. Social values often go beyond demand conditions.
5. *Technology* is transparent (both the existing conditions and possible future choices). It is as completely known and accessible to entrants as it is to the monopolist. The regulators also have complete knowledge of present and alternative future technologies.	5. Actual and alternative technologies are often complex and incompletely known to outsiders.
6. *Frictions in adjustments:* there are none. Changes are instantaneous.	6. Frictions and irreducible lags in adjustments are often large.
7. *Barriers to entry:* there are none (other than those created by scale economies or joint costs in the multiproduct case). Barriers to entry by low-cost firms can only be provided by actions of the regulators.	7. Barriers to entry are often large and have many sources.

8. *Competitive advantages of the monopolist* (in knowledge, financial ties, customer loyalty, experience, etc.): there are none.	8. The monopolist often has large advantages over entrants.
9. *Responses or pre-emptive actions by the monopolist:* there are none, and the entrant expects none to be made (by the Cournot-Nash assumptions).	9. The monopolist typically responds sharply to new entry, and entrants usually expect retaliation by price or other actions.
10. *X-efficiency and innovation:* the monopolist is expected to optimize them under perfect regulation.	10. Monopoly effects on innovation often remain. Also, the side effects of regulation itself often raise costs and enlarge investment (the cost-plus and rate-base effects).
II. "Dynamic" Analysis. In addition to the 10 static conditions:	
11. *Construction costs* are (a) large, relative to current costs and (b) have strong economies of scale (average cost declines steeply in the relevant range).	11. Construction costs vary but are usually below 20 percent of current costs. Scale economies are often absent or moderate.
12. *Construction occurs instantly,* at the time selected. There are no lags and no duration of construction.	12. Construction usually is lengthy, with a duration of two to five years.
13. *Strategic behavior is absent.* There is no possibility of pre-announcement, pre-emption, or other action designed to discourage entry or influence its timing. (Cournot-Nash assumptions are extended to this context.)	13. Strategic behavior can be sophisticated and strong.

8, the monopolist is completely exposed to entry, without advantages, frictions, or barriers. Nor by the Cournot-Nash assumption (assumption 9) is the monopolist expected to retaliate at all against entrants, or even to threaten to respond.

The results under regulation are assumed to be "ideal," maximizing consumer surplus. Because of the cost conditions, strict marginal-cost pricing will cause the firm to run financial deficits. Therefore, prices can do no better than to fit some second-best pattern. For the multiple-output case, BBWP propose Ramsey prices, a discriminatory multiproduct price structure. It is important to recognize that *Ramsey prices are virtually nothing more than price discrimination*. Prices are set short of full monopoly price discrimination by some uniform ratio. Such prices may well maximize consumer surplus, when (but only when) a deficit is otherwise inevitable.[15]

Ramsey prices may even arise voluntarily if a utility faces a moderate degree of entry, according to BBWP. A "weak invisible hand" may tend to result in Ramsey prices, by the monopolist's voluntary choice. (That apparently occurs only if no excess profits are permitted and if the costs of entry are high enough to deter superior potential entrants.) The outcome can maximize static consumer surplus.

If an entrant is permitted to produce a subset of products at lower than the Ramsey prices for those products, the utility's price vector is said by BBWP to be no longer "sustainable." The costs of the other products and overhead costs cannot be covered by the Ramsey prices. Therefore low-price entry can reduce consumer welfare, and it imperils the utility's existence. Still other harms have also been suggested, including "wasteful entry" and "destructive innovation."

A BBWP "dynamic" analysis also poses a two-period case where, under certain conditions, an entrant can construct new capacity (instantly) in the second period. Because the entrant's capacity can be built on a larger scale and therefore have a lower construction cost than the original firm's capacity, the entrant can displace the original firm altogether. Therefore competitive entry is said to threaten efficient production in both settings, static *or* dynamic.

Altogether, BBWP offer a seemingly powerful rationale for preventing competition in regulated sectors, particularly in the complex telecommunications sector.

Evaluation

In appraising this many-sided analysis, one begins with the assumptions. I will discuss them as a group and then consider several of them individually.

The main BBWP assumptions appear to be self-defeating: they contradict the reality that the analysis is intended to clarify. The right side of Table 2 summarizes actual conditions in regulated firms, based on the rich literature on these subjects. At the static level, assumptions 2, 4, and 10 are doubtful in most cases, and assumptions 1, 3, and 5 through 9 are radically inaccurate. The two-period assumptions 11 through 13 are also implausible.

BBWP have indicated privately on several occasions that their analysis is merely an "exercise" designed to give "insights," but not pretending to deal with real sectors. Certain assumptions (especially assumption 9) are needed to make the analysis tractable, they say, and are admittedly unrealistic. Yet the published results have been presented by BBWP as if they had strong relevance, and they are gaining acceptance as being definitive, in some quarters. Therefore one must judge the BBWP analysis's relation to reality. As Table 2 indicates and as I will discuss more thoroughly, it is slight.

In the single-output case, BBWP focus on natural monopoly conditions, where Q_I is less than MES. Setting price at marginal cost will cause a financial deficit, and so some second-best result is needed. New competition would reduce the established firm's output and force a higher price, perhaps putting the firm in deficit.

But if Q_I exceeds MES, so that there is room for competition, then sustainability is no longer in question. Prices can be at cost, while all firms—established and entering—are viable. In this general case, the BBWP analysis adds little to well-established conclusions about the single-output case.

As for multiple products, the sustainability issue arises when economies of scope prevail throughout. Then joint-production and overhead costs are extensive, and all outputs have marginal costs below average costs.

Yet if Ramsey (discriminatory) prices are not required, the BBWP analysis loses its force. That can occur if enough of the firm's outputs have rising marginal costs, which lie above average cost. Ex-

amples include peak-load outputs, which are common among utilities because of fluctuating demand. In other directions, too, marginal costs may be above average costs, for regulated firms typically provide a variety of services outside the central set of joint-cost products. There is often room for competition in many of the regulated firm's services.[16] Therefore, the need for comprehensive Ramsey prices is a peculiar special case among regulated firms.

Even where Ramsey prices *are* appropriate throughout, they will cease to be the welfare criterion for those outputs where the new entrant is able to set its price below the original firm's marginal cost. Then joint cost ceases to be joint, and it applies only to the remaining output, as noted above. The former monopolist may need to write off the joint cost and adopt a new set of prices. It is optimal to accept this entry, even if it renders the Ramsey prices unsustainable.

Sustainability therefore needs protection only (a) where the entrant's price falls in the band between the original firm's average and marginal cost, and (b) where no marginal costs are above average costs. This is a limited and perhaps small problem. Moreover, it is one of degree, not of absolutes. Entry may require adjusting the Ramsey prices, but perhaps by little. Any loss of consumer surplus may be slight. BBWP have neglected the question of degree, instead discussing only if sustainability exists or not. That has implied a more ominous effect of entry: that the firm's very existence is unsustainable. Why would BBWP give the problem such extended analysis if it did not have serious effects of that sort?

Yet the effect may instead be trivial: a robust established firm—dominant in its markets and with a wide range of pricing and strategic choices—merely adjusts one or several of its prices, or lets an entrant replace one of its products.

BBWP portray the monopoly firm as a passive target, unable to fend off entrants, even little ones. The Cournot-Nash assumption rules out actual responses, and the entrant expects to encounter none at all. This extreme assumption permits an analytical result to be derived, but it also leaves the analysis out of touch with reality. Instead, the monopolist is likely to retaliate—perhaps with great effectiveness—by using pricing and other strategies so as to penalize and deter entry. Such likelihoods are clear from the massive literature and experience in a variety of utility and industrial markets.[17]

Indeed, *the actual or threatened entry automatically shifts the monopoly's demand curves for the subset of affected outputs, making demand more elastic. Therefore it leads the monopoly to cut its price to meet—if not to undercut—the entrants' price, at least in the short run. If the utility looks ahead strategically, then it may cut prices—perhaps temporarily and strategically—enough to deter all anticipated entry. That action (or possibly just the threat of it) can often forestall any entry.*

In each case the actual conditions will govern the scope for deterrent actions and the probabilities of success. As a matter of general theory, the assumption of no retaliation is eccentric. As a matter of industrial experience, it is thoroughly unacceptable. One might similarly begin an aeronautical study by assuming that gravity makes objects fall upward. The Cournot-Nash assumption alone creates an important bias in virtually all of the BBWP analysis and its results.

The whole topic is one of degree as well as of logic. As for logic, the sustainability literature has first adopted a strange version of the regulated firm. It has treated competition as a threat to a narrowly defined price vector, suggesting that the firm's viability and existence will be in danger. But the effects of competition pose problems only if any likely social losses exceed the competitive improvements in efficiency, innovation, and other directions. Those matters of degree depend on the cost functions, demand conditions, and other relationships.

New competition may usually cause some degree of adjustment of relative prices, of a minor, rather than large or decisive, degree. The resulting efficiency losses, if any, may be trivial, sizable, or large, and they may be offset by other competitive benefits.

BBWP give us no guidance on that, nor does their analysis even permit us to weigh the quantitative effects or to ascertain whether core or adjacent goods are involved. This is the obvious line for serious research to take, but BBWP instead treat "sustainable" as merely a yes-no issue.

The same problems of unrealism and inattention to degree also afflict the "dynamic" analysis. Their analysis is significant only if construction costs are substantial relative to current costs, and if scale economies in construction are large. They also assume that construction can be completed instantly, and that no strategic actions about construction (e.g., pre-announcements, delays, etc.) are

possible. If either of the latter two assumptions fails, then the monopolist's ability to deter entry may be large.

If regulation is removed entirely, then two main possible cases arise: economies of scale and scope can guarantee the continuance of the firm at efficient prices; and if outputs are separable, then competition may spread, but the original firm can survive as much as its efficiency warrants.

In this light, the optimal choice of regulatory policy depends on (a) the static cost and competitive conditions (economies of scale and scope, separability, cross-over range, etc.), (b) the degree of X-inefficiency and retardation of innovation which may occur under monopoly power, and (c) the practical problems of imperfect knowledge, of the strategic advantages of the incumbent firm, and other such real factors.

One policy error therefore is to retain formal regulation after cost and demand conditions have shifted in ways that would give true separability and room for competition. Even if regulation is effective, the artificial monopoly is unnecessary. The monopoly is prone to have price discrimination, excess profits, X-efficiency, and slow innovation. Weak regulation will let these losses be bigger. A correct analysis will compare such losses with the possible distortion of prices that may be caused by excessive entry. That evaluation will need to allow for the practical problems of imperfect regulation, unequal knowledge, strategic moves, and the like.

The opposite error—to remove regulatory controls prematurely, before the cost conditions permit effective competition to force the monopolist's market share below the cross-over values—can also be costly. It gives the monopolist a free hand, while pretending that competition can restrain it. Such premature deregulation is the error that BBWP analysis encourages.

A key need is to set the burden of proof correctly. Experience suggests that the burden of proof should favor entry where cost and demand conditions probably permit M_0 to go below the cross-over zone. Willig's 1979 paper affirms that, for he shows that any remaining economies of scale and scope will permit the monopoly to deter inappropriate entry.

Meanwhile, events since 1976 have moved rapidly, permitting competition and new entry in a variety of regulated sectors. Indeed, the sustainability literature already has a certain antique quality. An

original stimulus was the Bell system's concern about new entry. But during 1979-80 Bell officials reversed their course to accept competitive policies, with little concern about the static niceties of sustainability. Apparently they are aware of their own competitive advantages, which the BBWP analysis has assumed away.

Summary

Because BBWP have tried to graft an analysis of competition onto an odd model of monopoly, their lessons are of limited relevance. A fuller analysis of the varieties of competition—especially quasi-monopoly and oligopoly—is needed. That can best be done by using models that embody the matters of degree.

The problem has elements—cost conditions, degrees of room for competition, cross-over zones, core and adjacent products, possibilities for strategic behavior, X-efficiency, and innovation—which usually occur in continuous functions. To be relevant to these issues, analysis needs to incorporate the continuity and make comparisons of degree. The BBWP papers have offered instead an either-or quality and little basis for evaluation.

By contrast, the familiar literature on dominant firms and oligopoly gives a basis for clarifying the move from monopoly toward competition. Variations in scale economies, of room for competition, strategic behavior, X-efficiency, and innovation can be estimated and compared usefully.

CONTESTABILITY: ANOTHER "NEW" ANALYSIS

More recently, the BBWP authors have shifted to a new approach, which Baumol (1982) presents as a "rebellion" which is a "common cause" shared by a large number of scholars.[18] It has been presented with much ballyhoo and self-praise. Though Baumol claims that it is relevant to a wide range of markets, its lessons—if they are valid—would immediately affect the deregulation of sectors such as telecommunications. In this brief space I can give only a summary evaluation of contestability. I will first note its general features and then assess the major lessons which the BBWP group draws.

Features of Contestability Analysis

First, the contestability analysis focuses on extreme conditions of free entry and absolutely vulnerable markets. "Entry is absolutely free, and exit is absolutely costless." Therefore: "The crucial feature of a contestable market is its vulnerability to hit-and-run entry" (Baumol, 1982). The analysis converts Bain's interest in entry conditions into an obsession. In this exotic world, entry-exit conditions supersede all others: if contestability exists, then it is immaterial whether the internal structure of the market is pure monopoly, dominant firm, tight oligopoly, or any other condition. If such an assumption is valid, then contestability is valid. If it is not, then contestability is not, or rather it is simply a portentous term which adds little to the mainstream analysis, in which entry is just one element that may have significant effects.

Baumol-Panzar-Willig (BPW) evidently hope to be able to prove that assumption, but the weight of the expert literature is that they cannot. As I noted above, the opposite is more likely; internal structure is central to the problem of monopoly, while potential entry is peripheral. Such a focus on internal condition does not reflect naive "structuralism," but rather many decades of thorough research by a wide variety of specialists. Only if internal structure is utterly featureless—with no inelasticity of demand for any consumer, no brand loyalties, no lags, no sales forces, no financial advantages or continuity, no possibilities of strategic action, etc.; in short, with none of the substance of real markets—might it be absolutely superseded by potential entry. In claiming generality, BPW neglect the nature of real markets, apparently willing to believe that an IBM or an AT&T can be instantly eliminated or neutralized by the possibility of "hit-and-run" entry. Such theorizing is not easy to take seriously as more than armchair thoughts. It will take an enormous fund of new industry studies, reversing the findings of decades of careful research, in order to make that claim believable. In short, the BPW analysis may be interesting theory, but it is not part of the field of industrial organization.

Second, apart from the one-dimensional analytical basis is the empirical question: *are* many important markets actually highly contestable? Air transport and trucking are two candidates, where the leasing of equipment makes possible relatively quick entry and exit.

There may be some others, but they are mostly in small, "light" industries, not in the leading familiar cases of market power, and certainly not in telecommunications.

Yet even the prime case of air transport (the focus of Bailey's interest) is turning out to be only moderately contestable. Its deregulation has led to rapid changes in many city-pair routes, and entry and exit are now virtually free. But many routes still have only one or two airlines. Pricing is fitting closely the predictions based on market shares; higher market shares cause higher price-cost margins, often by a multiple of cost. Even in this favorable case, "contestability" is either limited, or ineffective, or both. Air freight may be a slightly better case, but no serious competitor there can operate with leased equipment on a "hit-and-run" basis.

Trucking may fit the BPW analysis, but that remains to be seen. Other markets with extreme contestability—with possible "hit-and-run" entry, which is not only powerful but also more powerful than existing competition—appear likely to be few. They will include few, if any, of the familiar cases of major imperfect markets with firmly established leading firms.

The BPW group use perfect competition as a hapless foil in urging that contestability is radically new and better: a "benchmark for desirable industrial organization which is far more flexible and is applicable far more widely than the one that was available to us before." One could only say that either (a) strictly as theory or (b) by ignoring the mainstream of the industrial organization literature, which tries to combine structural and behavioral (and strategic) elements in a balanced study of general patterns and real markets. By comparison, contestability emerges as strictly a deductive analysis, whose far-fetched assumptions will not support extravagant claims. Unless subsequent work provides massive new one-sided evidence, contestability will join sustainability as a mere curiosum.

BPW offer several specific conclusions about contestability, which I will now consider briefly. All of them rest on the assumption of perfect contestability.

1. "In contestable markets, zero profits must characterize any equilibrium, even under monopoly and oligopoly." Therefore, "The analysis extends enormously the domain in which the invisible hand provides perfect efficiency." The

hand "seems to rule almost everywhere" that perfect contestability exists.

This is an ambitious claim, indeed. BPW here go beyond even the more extreme Chicagoans, for whom two firms in a market can be enough to guarantee optimality. *One* firm is enough for BPW. From sustainability, with its weak monopolist, we have come to monopoly which is competition.

Whether the claim should be taken seriously as more than a theoretical deduction depends on its analytical validity, as well as its factual strength. I have questioned those already above. BPW's conclusion looks like either a tautology derived from extreme assumptions about featureless markets and microsecond entry; or a bold, unrealistic assertion denying the richness of conditions in real markets.

2. Price will be driven toward marginal cost, neither less nor more. Therefore allocation will always be "first-best" efficient (except possibly for pure monopolists under certain odd conditions). As a side benefit, this is said to guarantee that no established firm (even a monopolist) will or can adopt predatory pricing. Apparently even the smallest resort to setting price below marginal cost will expose the monopoly firm to elimination.

   These lessons also appear to be tautologies arising from extreme assumptions. Any realistic imperfections in the market will undercut their validity. For example, lags or asymmetric knowledge alone can permit important deviations from the deduced long-run results. So will many other familiar conditions. As for predatory actions, the literature affirms that price-marginal cost inequalities are only a subset (perhaps a small one) of all truly effective predatory actions.

3. But BPW urges that even a pure, natural monopolist may be vulnerable to displacement by an entrant that, because it creates unnecessary new capacity, causes social waste. This is the intertemporal case noted above at Table 1 (p. 92). It deviates from the perfection provided by contestability; but only to stress even more flatly the impotence of existing monopoly.

> I have noted above that the assumptions underlying this point are not credible, and so it has no more validity under "contestability" than it did under "sustainability."

In summing up the material so far, there is little reason to alter the evolving mainstream analysis of industrial organization to accommodate a theorists' game resting on so narrow and shaky a base. So far, contestability and sustainability seem to be naive ideas with little substance. They are a diversion, but not a harmless one, for they aim to replace a solid research basis with theological speculations. These speculations seem to be of little relevance either to the main industries posing policy issues or to the design of those policies themselves. The exercise seems similar to saying that if mountains are as flat and as smooth as ice, you can skate across them effortlessly. Can a complicated analysis beginning with such bizarre assumptions be fruitful, if it displaces more realistic treatments?

One is left instead with the evolving set of prosaic structural and behavioral tools to deal with competitive issues and policies. In the case of telecommunications, one needs to consider the room for competition, the need for dominant-firm shares to go below about 50 percent in order to provide for effective competition, the need to provide access for new competitors and protection against predatory actions.

One feature of competitors' access needs special comment here, for it links recent concepts of entry deterrence with the critical problem of interconnection in the telephone sector. A monopolist can discourage entry by building excess capacity ahead of normal growth needs, as Alcoa and other monopolists have known and as Spence and others have recently analyzed. Excess capacity growth can even be effective in controlling competitors already in the market, as DuPont Co. recognized in the titanium dioxide market after 1971.

That incentive might induce a telephone monopolist to overinvest, so as to discourage new and existing competition. But such excess capacity can have a reverse effect, when competitors have access to share the use of that capacity. Then the greater the capacity, the stronger the case for admitting additional competitors; and the effi-

cient price to charge for access varies inversely with the degree of excess capacity.

This ambivalence in competitive effect may cause investment in such a common carrier to be indeterminate. The firm wishes to have competition-reducing excess capacity but also to prohibit access to it by competitors. There may be a special need to inspect and possibly to constrain certain capacity decisions, if competition is to be effective.

POLICY TOWARD THE SECTOR

In approaching telecommunications policies, the main lessons are that (a) changes in policy need to fit the economies of scale, and (b) transitions to competition require sharp drops in the monopolist's market share, rather than just the possibility that potential competition will become actual competition. The BBWP euphoria about potential competition needs to be replaced by a cautious withdrawal of regulation *only* where competition is established. Deregulation is a complex balancing process in a sequence of steps which takes time. One should not apply simple checkerboard economies, which merely assign markets among the pure extremes of natural monopoly and natural competition—and perfect entry.

For all their strengths, the antitrust agencies cannot handle the large, complex problems of price strategies in this sector. Antitrust is weakest in treating dominant-firm problems. Much of the staffing and skills of the FCC would need to be transferred to the Antitrust Division, and lengthy lawsuits (both public and private) would proceed in the courts rather than at the FCC. To that extent, the consent decree would relocate the problem—possibly in a less skillful process—rather than solve it.

I will discuss the main lines of the 1982 decree, in its final form: it presents several clear gaps and flaws.

Revising the Consent Decree

Some parts of the antitrust settlement will probably increase both competition and efficiency.

1. The separation of the Bell Operating Companies (BOCs) from Western Electric—perhaps the oldest antitrust objective—will encourage competition in BOC equipment purchases. Western Electric will also add competition in other markets for equipment.
2. Separate free-standing BOCs will now be exposed to direct market tests of their performance for the first time.
3. A separated Long Lines will provide a cleaner economic basis for determining the terms for competition in intercity service.

These changes therefore are improvements. Note also that they suggest that two frequently asserted kinds of economies within the Bell system are not as large as Bell officials have claimed in the past. One is vertical economies between Western Electric and the BOCs, some or all of which will now be "lost." The other is economies of scope among the parts of the unified network. *Apparently both economies are small, as many non-Bell observers have long suggested.* Or else the excess profits from unregulated pricing of the Long Lines and Western Electric goods are expected to be so large as to exceed the past economies.

Yet the decree has three basic flaws. First, it creates two parallel sets of monopolies: (a) the BOCs, and (b) the rest, mainly Long Lines, Bell Labs, and Western Electric. Second, it attempts to prevent future competition among them. Third, it leaves unanswered some highly complex problems of regulation. They will require sophisticated policy actions at least during the transition to competition.

By three simple revisions in the decree, these problems could be minimized by increasing the scope for competition.

The BOCs should not be limited from entering adjacent markets, including long-distance service, data transmission and processing, local programming, etc. Their local markets are likely to become at least partially competitive as cable technology develops. Perhaps more important, *the BOCs are a major natural source of effective competition for Long Lines.* Encouraging that competition, rather than suppressing it, may be the most important method for eventually placing Long Lines under effective competition so that deregulation can safely occur.

One neglected issue has been the proper size for the individual BOCs. AT&T officials are planning a grouping into seven regional units. That may render state regulation ineffective, by spreading each BOC over an average of seven states. But it might permit more effective BOC competition in intercity traffic. The issue needs careful study, with the most efficient solution probably lying between seven and fifteen new BOCs.

Long Lines will possess great market power in a highly profitable, rapidly growing market. The current competitors are small and highly vulnerable to Bell's pricing policies. They and other possible competitors could be easily excluded or limited by Long Lines (and the BOCs) after separation occurs. Therefore, any unregulated basis for Long Lines now seems naive, since competition could be readily eliminated or controlled. Either regulation will have to continue to an important extent, or Long Lines will have to yield down to a market share below the cross-over range, which is probably around 50 percent. Or some other basis for guaranteeing effective access and competition in intercity service must be provided.

It would be most efficient to separate Long Lines from Western Electric and Bell Labs. Any vertical economies are presumably small, as has been noted, especially compared to those between Bell Labs and the BOCs. Pooling Long Lines' large flow of profits with the equipment side would (a) permit exclusion in Long Lines' purchasing policies and (b) permit cross-subsidizing of Western Electric's competitive strategies in selling equipment to BOCs and others. Having partly removed this incubus by separating the BOCs from Western Electric, it is clearly best to remove it completely by making Western Electric entirely independent. Under the actual decree, large regulatory and antitrust problems will continue.

Further Problems of Regulation and Control

Now I turn to the problems that would continue even if the decree had fitted my suggestions—the problems of access, the treatment of Long Lines, and the future role of IBM.

If competition on intercity routes is to be viable and efficient, then the access to interconnection and to bulk use of Long Lines must be open and neutral. Such access may also be important in the

evaluation of an optimal set of local services, including telephone, cable, information services, and so on.

Currently, access is uneasily governed by legal and FCC administrative criteria, trying to offset the natural Bell tendency to set high prices for its competitors' access. The settlement will make access terms more indeterminate than they are already. Each local BOC will apparently be free to seek whatever terms it prefers, with jurisdiction perhaps held by the FCC, or perhaps by the state regulators, or both. There may be volume discounts or other selective pricing, plus various restrictive provisions, which could prevent fair competition.

Access is a core service which is like a public good. It is probably not neatly separable from other operations, but it needs nonetheless to be separated from the incentives of the BOCs and Long Lines. Outside regulatory efforts to control it will be ineffective, especially because technology is changing rather than fixed.

A separate publicly owned and operated entity to sell and control access may be the best solution. Its objective would be to provide access at neutral conditions, which (a) prevent any overcharging for access, and (b) prevent discriminatory prices among users. It would need to have thorough knowledge and to operate as an owner of the interconnect equipment as well as possibly some transmission capacity.

Being public and neutral would not guarantee efficient results. The correct cost bases would still need to be applied, often involving complex conditions. But it would avoid at least some of the awkwardness and indirectness of the present FCC attempts to control policies.

The treatment of Long Lines will be especially difficult, because it involves a virtual monopolist controlling complex and changing technology, under demand conditions permitting high excess profits. Entry by small firms since 1978 has already demonstrated the difficulty of reaching correct pricing levels for access. To get effective competition will probably require moving Long Lines' market share on all major lines below 50 percent, as noted earlier. Yet that will presumably eliminate Long Lines' role as being responsible for service availability. How to make that transition will be complex, given that the other newcomers will be vulnerable to BOC policies controlling access and to Long Lines' pricing and related actions.

During this transition these conditions will need to be regulated, at the very least, until fully effective competition is established, with Long Lines and BOCs unable to eliminate any efficient competitor or entrant. But it is not clear that such an efficient outcome can be reached. For one comparison, the transition now occurring toward full airlines competition is still very much in doubt; high degrees of monopoly remain in many submarkets and tacit collusion is not absent from numerous trunk routes. And in airlines the technology more clearly favors competition than it does in telephones. Also, airline and air cargo competition is possible partly because the crucial "interconnection capacity"—embodied in airports—is provided by public enterprise which is entirely separate from the airlines.

Telephones are perhaps more comparable to electric power. Competition in bulk power generation and transmission is feasible and could be effective in large areas. But the transition to that result would be (a) complex, (b) incomplete in many parts of the country (where users would have only one or two potential suppliers), and (c) perhaps inferior to moving instead to a unified transmission grid. Such a public grid could treat all parties neutrally in pricing and access. The choice among the status quo, unorganized free competition, or a public grid, plus competition, is an uncertain one.

Similarly, in telecommunications, it is not clear that open competition under transitional regulation is a superior treatment for Long Lines than is the status quo (under close regulation) or a public enterprise controlling access. Nor are the best forms and policies of such a public unit clear. It might be sufficient for it to be the leasing and contracting agent. Or the unit might also need to own satellite capacity and/or land-based facilities adequate to force the prices of Long Lines and other firms down to cost. Or the public unit might need to own and supply the interconnection technology in order to prevent Long Lines, the BOCs, or any other group from impairing free access.

At any rate, the problem of access poses difficulties far greater than any simple "unregulated, competitive" approach can solve. Both structure and behavior are involved.

Some observers foresee a major role for IBM in the sector; some have credited Mr. Baxter with ending both cases simultaneously in

order to bring AT&T and IBM into a duopoly situation. The two would then offset each other's dominance.

This idea unfortunately has little substance. Granted, there may be direct competition between IBM and AT&T in data processing and electronic equipment. Though this is a relatively minor overlap, the two firms might shift more directly into each other's product lines.

But at present that is only a hope. Such a shift would occur only if each firm believed that risk-adjusted profit opportunities in the other's markets exceed its own alternatives by more than the costs of developing the new products, after allowing for retaliation. Each firm can penalize the other; entry is not a simple hit-and-run situation but a strategic matter. Each firm has powerful sales forces, but only in its own product line.

Therefore there are large impediments to cross-entry and future competition. The two firms are likely to maintain their separate specializations—in which each is well ahead of the other—rather than to begin actions which will jointly minimize their profits.

This conclusion suggests that the contestability arguments may have little relevance even when the two potential entrants are powerful.

To summarize this section, the AT&T antitrust settlement will yield only certain real but limited gains. It would be more efficient to separate Long Lines also from Western Electric and Bell Labs. Long Lines will continue to need to be regulated until effective competition is established, possibly with the aid of a public entity to assure equal access.

The BOCs should probably be permitted to enter any business, including long-distance transmission and local programming activity on cable, while cable and other entities should also be permitted entry into BOC markets. Meanwhile, the proposed new BOCs are probably too few to permit effective state regulation.

Indeed, the entire settlement leaves much heavier, more complicated regulatory responsibilities than have yet been admitted. These must somehow be borne until competition advances enough to convert monopoly down into tight oligopoly or less. That shift may be slow or may never reach effective competition at all in major parts of the sector. Meanwhile, there need to be positive steps to promote

entry and protect new entrants. Far from ending regulatory and antitrust needs, the settlement will make them more complex.

CONCLUSION

I have offered four sets of ideas. First I noted the general background for events in this sector. Then I summarized the basic criteria of competitive markets, as developed in the mainstream literature on industrial organization. Next came several conditions important in telecommunications and similar systems, including room for competition, the cross-over range of effective competition, core services, joint costs, and responsibility for service. Then I assessed sustainability and contestability, both for their general value and for guiding policies toward this sector. Finally, I suggested several ways in which the antitrust settlement needs changing, and I pointed out problems that will need major policy attention in the future.

Here I will note only several points: mainstream concepts of competition offer relatively definite criteria for guiding a shift toward effective competition. Generally, market share will need to be below 50 percent and entry will need to be free.

Despite "sustainability" arguments, the established monopoly firms (regulated or unregulated) are not usually fragile and vulnerable to entry. Contestability appears to be an analysis based on extreme assumptions and conditions, and so it is apparently of little relevance to most markets, including those in the telecommunications sector.

The antitrust decree needs major revisions, including the separation of Long Lines from Western Electric and Bell Labs. It also needs to widen the BOCs' opportunities, especially as competitors in long-distance markets.

Even an optimal structural settlement will leave major regulatory tasks. Long Lines will hold high market power. BOCs will pose many interstate and intrastate issues of pricing and competitive access. A public enterprise entity may be needed to manage access to long-distance and interconnect facilities, in order to case problems of regulation and effective competition.

Antitrust policies will be needed to promote competition. Most telecommunications markets will not move from existing monopoly to effective competition without great care in assuring neutral conditions.

Rather than ending antitrust's and regulation's roles in most of the sector, the coming structural changes will only make them more complex and important.

FOOTNOTES

I wish to thank Harry Trebing, Ralph Bradburd, Eli Noam, David Sappington, Glenn Loury, John Panzar, and participants at the Arden House conference, for discussion of many of the issues in this article.

1. This rise and an explanation of the methods for measuring it are presented in Shepherd, W.G., "Causes of Increased Competition in the U.S. Economy, 1939-1980," *Review of Economics and Statistics*, forthcoming 1983.

2. The role of antitrust and other causes is assessed in Shepherd (1983), *op. cit.*

3. Conditions during and since the 1960s are assessed in Scherer, F.M., *Industrial Market Structure and Economic Performance*, Chicago: Rand McNally, 1980, revised edition; W.G. Shepherd, *The Economics of Industrial Organization*, Englewood Cliffs: Prentice-Hall, 1979; and W.G. Shepherd, "Economies of Scale and Monopoly Profits," in Craven, J., ed., *Industrial Organization Research*, Boston: Martinus Nijhof, 1982.

4. They are presented fully *inter alia* in Scherer, *op. cit.*, and Shepherd (1979), *op. cit.*

5. This is clear from results summarized in Scherer, *op. cit.*, and Shepherd (1979), *op. cit.* Shepherd (1982), *op. cit.*, discusses and estimates "surplus market share" in certain large U.S. firms.

6. See Kahn, A.E., *The Economics of Regulation*, New York: John Wiley & Sons, 1971, 2 vols.; Shepherd, W.G. and Wilcox, C., *Public Policies Toward Business*, Homewood: Irwin, 1979, 6th edition; Nelson, J.R., *Marginal-Cost Pricing in Practice*, Englewood Cliffs: Prentice-Hall, 1964; Bonbright, J.C., *Principles of Public Utility Rates*, New York: Columbia University Press, 1962; and Scherer, *op. cit.*

7. On price discrimination, see Robinson, J., *The Economics of Imperfect Competition*, London: MacMillan, 1933; Machlup, F., *The Political Economy of Monopoly*, Baltimore: Johns Hopkins University Press, 1952; Kahn, *op. cit.*; and Scherer, *op. cit.*

8. See Shepherd and Wilcox, *op. cit.*; Shepherd, W.G., *The*

Treatment of Market Power: Antitrust, Regulation and Public Enterprise, New York: Columbia University Press, 1975; and Nelson, *op. cit.*

9. The issues are reviewed in Scherer, *op. cit.*; Shepherd (1979 and 1975), *op. cit.* Dominant firms (with M_0 above 50 percent) usually diverge sharply from competitive conditions. There is no sharp borderline, and other factors (such as entry barriers) can affect monopoly power. Yet market shares and market power are positively related.

10. Occasional strikes, overloads and weather catastrophes are reluctantly accepted by regulators and customers, despite the utility's formal absolute obligation (as in New York telephone service during 1969-70 and the blackouts of 1965 and 1979). Regulatory penalties are applied carefully, often lightly and forgivingly. Franchises are rarely revoked, even for inept utilities. See Kahn, *op. cit.*; and Bonbright, *op. cit.*

11. The main papers are Panzar, J.C., and Willig, R.D., "Free Entry and the Sustainability of Natural Monopoly," *Bell Journal of Economics 67*, 1977, pp. 1-22; Baumol, W.J., Bailey, E.E., and Willig, R.D., "Weak Invisible Hand Theorems on the Sustainability of Prices in a Multiproduct Natural Monopoly," *American Economic Review 67*, 1977, pp. 360-365; Panzar, J.C., and Willig, R.D., "Economies of Scale in Multi-Output Production," *Quarterly Journal of Economics*, 1977, pp. 481-494; Baumol, W.J., "On the Proper Cost Tests for Natural Monopoly in a Multiproduct Industry," *American Economic Review 67*, 1977, pp. 811-822; Willig, R.D., "Multiproduct Technology and Market Structure," *American Economic Review 69*, 1979, pp. 346-351; Willig, R.D., and Baumol, W.J., "Intertemporal Unsustainability," manuscript, 1980; Bailey, E.E., "Contestability and the Design of Regulatory and Antitrust Policy," manuscript, 1980; and Baumol, W.J., Panzar, J.C., and Willig, R.D., *Contestable Markets and the Theory of Industry Structure*, San Diego: Harcourt Brace Jovanovich, 1982.

12. The only previous published review of sustainability has been by Snowberger, V.C., "Sustainability Theory: Its Implications for Governmental Restriction of Entry," *Quarterly Review of Economics and Business 18*, 1978, pp. 81-89; which is a brief summary of Snowberger, V.C., "Efficiency, Sustainability, and Ramsey-Optimality of a Multi-Product Regulated Monopoly," Contractor's Re-

port, Office of Telecommunications Technology, Executive Office of the President (September, 1977). It deals only with the earlier papers.

13. That literature includes a series of Chicago-school critiques inspired by George J. Stigler, another series of studies sponsored by the Brookings Institution, arranged (and some of it written) by Harry M. Trebing at his Institute of Public Utilities at Michigan State University.

14. Frank Ramsey's 1927 note on taxation provided a simple rule for minimizing the impact of taxes. The BBWP analysis has adapted this rule for the declining-cost, inevitable-deficit case of a public utility. See Ramsey, F., "A Contribution to the Theory of Taxation," *Economic Journal 37*, 1927; pp. 47-61.

15. Static consumer surplus is not a complete criterion of social welfare. It may not be a well-defined maximand. It ignores X-efficiency and innovation. Also, there may be other social values not included in consumer surplus. Yet BBWP analysis considers it exclusively.

16. Such mixed sectors are discussed in Kahn, *op. cit.*; Bonbright, *op. cit.*; Shepherd (1975), *op. cit.*; and Phillips, A., ed., *Promoting Competition in Regulated Markets*, Washington, D.C.: Brookings Institution, 1975.

17. The literature on the complex conditions of entry has grown to a large volume. See Bain, J.S., *Barriers to New Competition*, Cambridge: Harvard University Press, 1956; Scherer, *op. cit.*; Klass, M.W., and Shepherd, W.G., eds., *Regulation and Entry*, East Lansing: Michigan State University Institute of Public Utilities, 1976; Shepherd, W.G., "Entry as a Substitute for Regulation," *American Economic Review 63*, 1973, pp. 98-105; Kahn, *op. cit.*; and Phillips, *op. cit.*

18. Baumol announces it in his AFA Presidential Address as follows:

> I believe it offers a host of new analytical methods, new tasks for empirical research, and new results. It permits reexamination of the domain of the invisible hand, yields contributions to the theory of oligopoly, provides a standard for policy that is far broader and more widely applicable than that of perfect competition, and leads to a theory that analyzes the determination of industry structure endoge-

> nously and simultaneously with the analysis of the other variables more traditionally treated in the theory of the firm and the industry. It aspires to provide no less than a unifying theory as a foundation for the analysis of industrial organization. I will perhaps be excused for feeling that this was an ambitious undertaking.

His article refers to a book-length treatment (Baumol, Panzar and Willig, *op. cit.*) which was unavailable at this writing. My comments therefore are necessarily provisional.