Regulation and Entry
Energy, Communications, and Banking

Edited by
Michael W. Klass
and
William G. Shepherd

1976
MSU Public Utilities Papers

Division of Research
Graduate School of Business Administration
Michigan State University
East Lansing, Michigan
Contents

Chapter I. Introduction
MICHAEL W. Klass
Senior Research Staff, Charles River Associates
WILLIAM G. SHEPHERD
Professor of Economics, University of Michigan

Chapter II. Entry, Regulation, and Efficiency
MICHAEL W. Klass

Chapter III. General Conditions of Entry
WILLIAM G. SHEPHERD

Chapter IV. Entry into Power Markets
JAMES R. NELSON
Professor of Economics, Amherst College

Chapter V. Entry Conditions in Telecommunications
HARRY M. TREBING
Professor of Economics and
Director, Institute of Public Utilities
Michigan State University
WILLIAM H. MELODY
Associate Professor of Communications and Economics
Annenberg School of Communications
University of Pennsylvania
I. Introduction

Michael W. Klass
and
William G. Shepherd

Economic analysis of regulation has advanced considerably since 1960, with increasing skepticism that regulation is beneficial. The case for abolishing or reducing regulation has been made—quite specifically for several sectors and commissions, most notably natural gas, railroads, and banking. Broadly speaking, the scholarly burden of proof has now shifted against regulation as an economic tool. Meanwhile, actual regulation has tended more to spread and harden than to diminish. The topic is complex, involving industry life cycles and the evaluation of regulation itself over time.

This book addresses the main practical form of deregulation, namely, new entry into "regulated utility" markets. If "potential entry" is as critical a factor as has been asserted in the recent literature of industrial organization, perhaps it should become a central

---


Michael Klass and William Shepherd

In chapter IV James R. Nelson analyzes the electricity sector. He resolves it into its three levels (generation, transmission, and distribution) and focuses on the role of vertical integration. Telecommunications is surveyed by Harry M. Trebing and William H. Melody in chapter V. They give special attention to the recent moves to permit competition in some submarkets (terminal equipment and bulk transmission), and the authors draw fairly definite lessons. In chapter VI Thomas G. Gies treats entry in commercial banking, drawing on his experience with attempts to increase banking competition and performance.

The authors bring up much direct experience and research in their topics, and all of them are open-minded, even hopeful, about entry and its possibilities. In chapter VII a shared theme emerges about entry and deregulation, but these are meant only as first, not last, words on the subject.

Note: This volume is one tangible result of the Workshop on Regulation and Potential Competition at the Department of Economics, University of Michigan. The workshop was supported by a grant from the Brookings Institution under its Program on Studies in the Regulation of Economic Activity. We wish to record our gratitude for the grant and for the patience and support of Joseph A. Pechman and Barger G. Noll at Brookings.
II. Entry, Regulation, and Efficiency

Michael W. Klass

The perennially horrified discussions of the prospects for improvement in economic regulation assign great significance to the effects of regulatory restrictions on entry and exit. Relaxation of these restrictions is to restrain (if not neutralize) any monopoly power of the newly deregulated firms, enforce efficiency, and encourage innovation. This chapter explores the interacting effects of entry restrictions on the behavior of regulated firms, potential entrants into the regulated industries, and the regulators themselves. The focus will be on the possible effects of entry restrictions and their relaxation on pricing and efficiency. The first section analyzes various proposed rationales for entry regulation; the second discusses the relationships among the form of regulation, barriers to entry into regulated fields, and the incentives to enter them; the third considers the effects of entry and its regulation on efficiency and innovation, emphasizing the possible relationships of entry and regulatory lag; and the fourth section offers conclusions.

What Is Regulated and Why

Control of entry into regulated markets generally proceeds via requirements that producers be licensed or certified by the appropriate authority. This control may go well beyond simply allowing or
disallowing entry into some or less broadly defined market. For example, in the trunk airline industry CAB regulation not only has limited the number of lines, but also attempts to control the size distribution of firms in particular city-pair markets. In the areas of electric power and telephone communications, franchises or licenses define the geographic area of operation of the regulated firms. Combined with the protective attitude of regulators, the effect is sharply to restrict potential competition from firms in adjacent areas. Regulatory restrictions on entry apply not only to potential de novo entrants into the regulated industries, but also to entry or entry equivalent behavior by firms already serving these markets.

Entry Regulation as a Redistributive Device

The rationales for the pervasive regulatory concern with entry range widely, some maintaining that entry restrictions are contrary to the public interest, being designed to benefit the regulated firms, others arguing that entry controls are necessary to assure efficiency. The first position is typified by George Stigler, who argues that regulation is demanded by the regulated firms, which rationally utilize the political process to increase their profits. In this view price regulation may arrange both the maintenance of cartel discipline and the division of profits among current and new producers. Entry controls are an integral part of an attempt to facilitate monopolization and, from the standpoint of efficiency, are to be condemned.

In a related view, put forth by R. A. Posner, the purpose of entry regulation is to preserve the profits generated by monopolistic pricing. These profits, instead of accruing to the owners of the regulated firm, serve to cross-subsidize services which, if provided at cost, would be forthcoming in smaller quantities. These services being pricing is area dce. Suppose that, with entry allowed, the effectiveness of entry is lowered, so that price is P, and quantity Q. The welfare loss is then area dce which clearly exceeds dce. By rendering regulation less effective, a more atomistic structure might lead to increased inefficiency.

Counterarguments are of course available. For example, the pre-

\[\text{de novo} \]

Completely neutral tax. There being no examples of such taxes, the case against cross-subsidization is difficult to make in general. Furthermore, granting the legitimacy of cross-subsidization, the entry controls needed to sustain the subsidy also must be regarded as legitimate.

The Stigler and Posner arguments might be characterized as explaining the existence of entry controls, yet not truly serving to legitimize them; the controls may still be attributed to serving the special interests of the regulated firms or some of their customers. Another type of argument begins from the premise that the purpose of regulation is to serve the public interest by promoting economic efficiency, and then argues that entry controls are needed to achieve this end.

Perhaps the most plausible argument of this sort is outlined below. Suppose that scale economies in an industry are such that only a few efficient firms could exist. Regulation might be advocated to prevent or moderate tendencies to monopolistic pricing which could result from recognized interdependence of firms. The regulators have limited resources to use in their investigations of costs and returns, and they are likely to function more effectively the smaller the number of firms in a market. Regulation of an oligopoly, with individual firms presenting differing and perhaps contradictory evidence on costs and demands, should be more difficult, more costly, and presumably less effective than regulation of a monopoly. The argument is presented graphically in Figure 1, in which D is industry demand, and MC is industry marginal cost, assumed independent of the number of firms in the industry.}

In a monopoly, industry regulation might succeed in establishing price P, and quantity Q. The welfare loss due to this monopoly pricing is area dce. Suppose that, with entry allowed, the effectiveness of entry is lowered, so that price is P, and quantity Q. The welfare loss is then area dce which clearly exceeds dce. By rendering regulation less effective, a more atomistic structure might lead to increased inefficiency.

Counterarguments are of course available. For example, the pre-
Input versus Output Efficiency

While it is plausible to argue that the regulators of a monopoly would be more effective in setting price close to observed cost than would the regulators of an oligopoly, the efficiency conclusions are more complex than the above discussion has indicated. Efficiency of the regulated industry has at least two dimensions. The first is output efficiency: Does the industry produce the output at which price equals the marginal cost of production? The second is input efficiency: Are input combinations chosen to minimize costs for whatever output is produced? Figure 1 addresses only the question of output efficiency. By assuming that MC remained fixed as price varied, the question of the effects of regulation on input efficiency and the level of MC were ignored.

The effect of regulation on input efficiency has been discussed extensively, beginning with the contributions of Harvey Averch and J. L. Johnson and S. H. Weitz. Application of some of the results of this literature indicates that the "facilitation of regulation" case for entry controls is dubious. To address this argument, some notation must be defined, following Averch and Johnson: 4

- $\pi$ = the regulated firm's total profit;
- $Z$ = the firm's output;
- $X_1$ = the quantity of capital in the rate base;
- $X_2$ = the quantity of labor input;
- $r_c$ = the market cost of capital per dollar of capital;
- $r_s$ = the wage rate;
- $r$ = the allowed rate of return;
- $Z = Z(X_1, X_2)$ = the production function, assuming diminishing marginal products; and
- $P = P(Z)$ = the firm's inverse demand function.

In this model it is assumed that the firm has complete knowledge of production and demand conditions. The allowed or fair rate of return, $r$, is assumed to be at least as large as $r_c$, the cost of capital, but less than the unconstrained profit maximizing rate of return, $r_s$, for this stage of the discussion assume that there is no regulatory lag: regulation is continuous, and prices are adjusted so that at all times

---

1Another dimension might be pricing structure efficiency: Is there price discrimination which results in inefficient allocation among consumers of the amount produced?

2The notation is that used by Baumol and Kleinwick.
the net rate of return is $s$. The firm is assumed to choose $X_i$ and $X_j$ to maximize total profit,

$$\pi = PZ - r_iX_i - r_jX_j,$$

subject to the fair rate of return constraint,

$$PZ - r_X X_i = s. \quad (1)$$

A standard result is that the imposition of the rate of return constraint, (2), results in input inefficiency in the sense that the regulated firm chooses a capital-labor ratio $(X_i/X_j)$ exceeding that which minimizes cost for the chosen level of output. Furthermore, the capital-labor ratio and, hence, the regulated firm's cost, depend on $s$. These points may be seen in Figure 2, which follows E. E. Zajac's

![Figure 2: Input Choice Bias in the Averch-Johnson Model]

Michael Klass

The curves $S_1$, $S_2$, and $S_3$ are iso-rate of return contours of a profit surface in $X_i - X_j - \pi$ space. Called constraint curves, they correspond to various levels of $s$, where $s_1 < s_2 < s_3$. If $s_1$ is the allowed rate of return, any capital-labor combination in the interior of $S_1$ yields a rate of return exceeding $s_1$ and so is forbidden to the firm. All capital-labor combinations on $S_1$ yield a rate of return of exactly $s_1$ and so are allowed. Since the amount of profit allowed is proportional to $X_i$, the rate base, the profit maximizing input choice for any level of $s$ is that point on the appropriate constraint curve with maximal capital. If $s = s_1$, the input choice is given by point $a$. The locus of all such constrained profit maxima is given by $(X_i/X_j) = \{X_i(X_j)\}$, the firm's expansion path under the Averch-Johnson regulatory constraint. $(X_i/X_j)$ is the firm's unconstrained expansion path, the locus of minimum cost input combinations. If the production function is homogeneous, $(X_i/X_j)$ will be a straight line through the origin (as in Figure 2), and efficiency dictates a constant capital-labor ratio. For $r_1 < s < r_2$, $(X_i/X_j)$ lies below the efficient path. If a typical isoquant is given by $Z$, the efficient input combination will be less capital intensive than the input combination chosen by the regulated firm.

Unfortunately, the direction of response of the capital-labor ratio to changes in $s$ depends upon the form of demand and production functions. W. J. Baumol and A. K. Kleiner have shown that if capital and labor are substitutes in the production of revenue, a movement of $s$ toward the cost of capital will raise the capital-labor ratio — that is, $d(X_i/X_j)/ds > 0$. Thus, an improvement in the performance of the regulators, in the sense of moving the average price of output closer to the firm's average cost of production, may well reduce efficiency by raising the level of these costs. The resulting resource waste could easily exceed any possible gains in terms of output efficiency. Alternately, a fall in $s$ may induce input choice bias, lowering costs and increasing output. In such cases, the efficiency effects of entry restrictions which allow more effective regulation would be unambiguously positive.

*See Baumol and Kleiner, pp. 166 – 67, or Zajac.

*This result holds in general and is demonstrated by Baumol and Kleiner, p. 175.

*See Baumol and Kleiner, p. 179.
Another efficiency oriented rationale for entry restrictions would apply only to situations in which firms have decreasing average costs. In such cases, price discrimination may be used as a mechanism to increase output toward the level at which marginal cost intersects demand and at the same time over total costs. Relaxation of entry restrictions might result in entry into the high profitability segments of the business, reducing revenues below costs and forcing decreases in output below the efficient level. If the decreasing cost regulated industries, such as electric power distribution, actually followed such pricing schemes, entry prohibition might have some potential efficiency rationale. The relative absence of such attempts at marginal cost pricing weights against the applicability of this argument.5

5See Shepherd.

Since $P'(Z) < 0$, the residual demand will be more price responsive than the initial demand function, the price elasticity of demand will be greater at any output, and so the desired price should be lower.6 If the price ultimately set is that which is optimal for the core firm, industry output should rise. The fact that firms wished to enter indicated that at prior prices consumers were willing to pay for more output.7 If the entering firms had the same costs as the initial monopolist, output efficiency would be improved. Barring the possibility that input efficiency would be worsened by a decrease in the attainable rate of return, overall efficiency would be served.

Suppose that the entrants suffer a cost disadvantage. The fact that they are willing to enter indicates that output efficiency will be served. And, although the cost minimizing method of obtaining the extra output might, especially in decreasing cost cases, have been for the initial monopolist to increase production, this option may not have been available. Unless forced to do so, a monopolist doing as well as he can under the regulatory constraint will have no incentive to increase output. Unless regulation functions to maintain industry output, entry, even by higher cost rivals is likely to serve efficiency, albeit in a second-best manner.

While, given the imperfection in regulatory processes, there is reason for concern. Suppose that a higher cost rival receives the blessing of the regulatory authorities and enters. The profits of the established firms will fall, and the likely response of the authority will be to raise prices, further aggravating the output distortion. Similar results could occur if the entrants had a cost advantage. Feeling protective toward the established firms, the authority may allow price increases, leading to inefficiency which could outweigh any cost savings.

Problems of Interdependence of Services

In some cases, interdependence between the availability of services supplied by a regulated firm and the demand for services offered by potential entrants can pose complex dilemmas for regulatory authorities. In the Avery-Johnson model this would be arrived at indirectly, through the firm's variation of its output mix. Whether this result, holding for an unconstrained monopolist, holds as well for a constrained firm is open to doubt.8 If the initial monopolist and potential entrants had increasing costs, excess profits could coexist with an output at which price equals marginal cost. Still, efficiency could be served by entry, which would result in smaller output and lower cost by the initial monopolist.

6In the Avery-Johnson model this would be arrived at indirectly, through the firm's variation of its output mix. Whether this result, holding for an unconstrained monopolist, holds as well for a constrained firm is open to doubt.
7If the initial monopolist and potential entrants had increasing costs, excess profits could coexist with an output at which price equals marginal cost. Still, efficiency could be served by entry, which would result in smaller output and lower cost by the initial monopolist.
authorizes. For example, MCI found a substantial demand for low priced and relatively low reliability microwave communications, and eventually it was given permission to enter one of Bell Telephone's markets. Yet, the level of this demand may have been overestimated on the availability, at current prices, of Bell's highly reliable backup service. If Bell is not allowed to respond to MCI's entry (and, presumably, to the entry of other such firms) by cutting price to the level of its own marginal costs, it may be motivated to reduce service levels or quality. Thus, in turn, could destroy the viability of the entrants, resulting in the wasting of the resources embodied in their fixed capital and proprietary knowledge. If Bell is allowed to respond, its profits probably will fall, facing the regulatory authority with the likelihood that some of the activities currently cross-subsidized out of Bell's microwave profits must be curtailed or, possibly, their prices increased. Yet, the open entry policy might still be preferred in that it could pressure established firms to expand the variety of their services and force commission to evaluate the wisdom of their cross-subsidization policies.

Do Entry Restrictions Matter?

Assertions that the relaxation of regulatory bar to new entry will result in competitive pricing are too sweeping. For this to be true there must, first, be no enduring nonregulatory barriers to entry into regulated industries. If there were, the regulatory restrictions might be redundant. But suppose the restrictions were not redundant, and that scale economies were significant and dictated firm sizes implying substantial market shares. Then the best that could be hoped for from unrestricted entry would be average cost pricing, with output still restricted relative to the competitive ideal. Whether overall efficiency is improved is then a question of the efficiency of regulation as compared to the efficiency of entry in pushing output toward the efficient level. And if regulation has encouraged discriminatory pricing which allows expansion of output toward the efficient level, unrestricted entry could undermine this situation and decrease efficiency. This would be a case of harmful cream skimming.

In this area, as in others, the effects of entry restrictions cannot be considered apart from the nature of the regulatory process. Harold Demsetz has argued that all that is required for competitive pricing is that there be a substantial number of competing potential providers of service and that the licensing authority grant the franchise to the firm offering to supply market demand at the lowest price. The form and content of the contract between the firm and the regulators are not made explicit, but evidently there is an agreement to sell to all customers at one agreed-upon price. Demsetz argues that even in supposed natural monopoly cases the competing bidders will undercut each other so long as price exceeds cost. The end result is supposed to be a zero profit solution which is also efficient. There are several problems with this argument. First, as L. D. Telsor has shown, while a zero profit solution may be reached, efficiency is not achieved. This situation is illustrated in Figure 3. If the contract between bidders and regulators specifies the meeting of market demand at a single price, any bid above average cost will be undercut. The winning bid will offer to supply quantity \( Q_c \) at
price $P_c$. Clearly, since $P_c = AC$, this is a zero profit solution. Yet, as Telszer has pointed out, this is not efficient in the usual sense of price equaling marginal cost. With demand and cost as shown, the efficient quantity and price would be $Q$ and $P$, respectively. Yet no firm would have an incentive to make such an offer.

For Demsetz’s bidding process to yield an efficient solution in the decreasing cost case, the contract must specify more than a single price for the corresponding market clearing quantity. It must either be an "all or nothing" contract for $Q$, or a contract requiring that the firm follow a discriminatory pricing policy, with the lowest marginal price being $P_c$. For regulators to follow such plans they must have extensive knowledge of both demand and marginal cost functions.

The Demsetz analysis is useful, not for its conclusion that some hypothetical bidding process can lead to efficiency, but because it points out that a firm’s acceptance of a franchise entails, in effect, a contract with regulators to perform specified tasks. The way in which the contract is let can influence the goodness of performance of the regulated firm. The analysis indicates that, even in natural monopoly cases, potential monopoly profits and power might be mitigated by allowing alternative potential suppliers to bid for the chance to provide service. Such processes could be similar in effect to take-over attempts.

Suppose that such an auction process yielded a zero profit solution for some period of time. The process would have to be repeated frequently, as demand and cost shifted. And while, de novo, there might be a competitively large number of potential suppliers, each with equal ability to produce, the situation would surely be different after one firm had served the market for some time. This firm would have learned to adapt to the peculiarities of the market, would be familiar with customers, and perhaps most important, would be accustomed to dealing with the regulators. Assuming that it remains efficient, it would have a cost advantage over other bidders and would be able to command a rent equal to the extra total costs of the next most efficient bidder. The moral might be to open up franchises to occasional attack and to use price and cost as elements in the choice of the winner. While perfect efficiency is not gained, a valuable constraint is placed on X = efficiency and slack, and possible monopoly power is limited.

The above arguments also point out that continuous, free, and open competition including unrestricted entry may not be necessary, in some cases, to achieve some of the real benefits of competition.

Michael Klass

Barriers, Incentives to Enter, and the Form of Regulation

The effects of entry restrictions must be evaluated by comparison with the situation that would prevail in their absence. Entry might be blocked: In each market segment the price prevailing in the presence of regulatory restrictions on entry might be below the price at which, absent the restrictions, significant entry would occur. In such cases it is tempting to argue that entry prohibitions have no effect and that their relaxation would have no moment. This view is incorrect for several reasons, which will be discussed in turn.

Entry Restrictions and the Quiet Life

First, it is possible that the entry prohibitions themselves would directly affect the behavior of regulated firms. For example, knowledge that they are free from challenge either from conventional entry by new firms or from take-over threats by other management groups could breed cost inefficiency in addition to any waste resulting from Averch-Johnson biases. Similarly, the sense of security could lead to relative indifference to cost reducing innovations. These effects would be greatest when regulation was most effective, for when regulators are able continually to hold the rate of return close to the cost of capital, the profit rewards to efficiency and innovation will be low.

Entry Restrictions and Nonregulatory Entry Barriers

The height of the nonregulatory barrier to entry is not independent of regulation. Major elements of nonregulatory barriers include scale economics and lack of technical knowledge on the part of potential entrants. While the extent of scale economies is largely determined by technical possibilities, legal prohibitions on entry render potential entrants less likely to produce or acquire the technical and economic knowledge requisite for effective competition. This argument indicates that, especially in the regulated sectors, the most likely entrants into a market area are newly formed firms, but current (regulated) operators of other market areas. These firms already possess much of the necessary knowledge and do not have to take the leap of deciding to acquire it in the face of a small probability of success. Effective entry via changes in management through merger could yield efficiency benefits without raising the problem of
increases in costs due to duplication of facilities or decreases in output of decreasing cost firms.

The presence of legal restrictions on entry means that a potential producer must incur not only the additional costs of small initial scale and experience, but also the legal costs of fighting the prohibition. This additional deterrence might be socially acceptable if one of the goals of entry regulation is to maintain an industrial structure in which detailed regulation is feasible.

Entry Restrictions and Capital Cost Barriers to Entry

The height of barriers to entry into regulated fields depends upon the regulatory rules in yet another way. Firms subject to rate of return regulation and protected by entry restrictions may be safer than unregulated firms and may thus enjoy a lower cost of capital. Following Stewart Meyers, two polar cases of regulation are considered. The first fixes revenues ex ante. In each period prices are adjusted so that revenues cover some prior estimate of operating costs plus the cost of capital. In this case investors bear all of the uncertainty of operating costs and also reap the benefits of any improvements in efficiency. The second case adjusts price so that in each period the realized rate of return equals the cost of capital or the return which investors expected to receive. Under this rule variations in operating costs or the cost of capital are passed on to consumers. Any risk borne by investors results from regulatory lag or the possibility that there may be no price which will yield the needed revenue. This attempt to equate actual revenues and costs ex post seems consistent with the regulatory process of the Averch-Johnson model. It is also consistent with the pervasive regulatory concern that investors be duly compensated. Investors bear no risk under this ex post type of regulation, and so the cost of capital should be lower than under ex ante regulation.

This being the case, it is possible that regulated firms face lower capital costs than typical unregulated firms. In the unregulated sector, at least some of any cost uncertainty falls on investors. This

---

Michael Klass

---

cost advantage could serve to increase the already high barriers surrounding many utility markets. Furthermore, the control of entry will facilitate the ex post type of regulation by maintaining the level and inelasticity of demand faced by the regulated firms. This possible advantage of regulated firms is yet another reason to look for potential entrants into regulated markets among the population of regulated firms.

Knowledge of the Regulatory Process as a Source of Scale Economies and Barriers

There are also powerful noneconomic, or partially noneconomic, reasons to expect entrants into regulated markets to be other regulated firms. To succeed in a regulated market, a management not only must know how to produce, but also must be able to deal effectively with regulatory commissions. The needed inputs include knowledge of the regulatory process and the ability to influence, both personally and politically, the decisions of the regulators. To some extent these resources can be purchased on an open market from specialized consulting and law firms. Yet, it may pay a regulated utility to employ persons with substantial regulatory influence, such as former regulators or former political officials. The presence of these indivisible resources could be a source of decreasing costs of dealing with commissions and could be a powerful incentive for mergers of regulated firms.

Averch-Johnson Effects and Incentives to Enter

The likelihood of Averch-Johnson bias and other forms of cost inefficiency complicates the question of the incentives to enter a regulated field. If the allowed rate of return exceeds the cost of capital, the costs of regulated firms will be above the minimum attainable. Other firms may be able to effectively enter the regulated market by producing close substitutes, yet escape the full force of regulation. Then even if the excess rate of return of the regulated

---

13One might ask whether this decrease in capital cost is a social or merely a private cost advantage. Note that risk is not eliminated by the regulatory rule — only its allocation among various potential risk bearers is changed. This, in itself, must increase welfare. Stewart Meyers (op. cit.) argues that the net extent of regulation may entail a high cost of capital, if this leads to a good allocation of risk-bearing. The relevant principle seems to be that risks should be borne by those economic units specialized in this function. This means investors rather than consumers. Capital markets afford a wide variety of alternatives for diversification and for tailoring of portfolios to particular preferences.
firms was very low, there could be substantial incentives to enter. This entry might well be by firms with costs above the minimum attainable by the regulated firms and so be efficient only if the effects of regulation are not as given. In fact, in such situations, entry may worsen cost efficiency.

Figure 4 depicts the supposed situation. Initially all production is by a regulated, decreasing cost monopolist facing the market demand curve DD. The monopolist’s minimum attainable or efficient average cost and marginal cost functions are $AC_m$ and $MC$, respectively. The firm is regulated, with the allowed rate of return, $s$, exceeding its cost of capital. The presence of an Averch-Johnson bias is indicated by the fact that the firm’s observed average cost, $AC_o$, exceeds $AC$. Assume that price and quantity under regulation are $P_r$ and $Q_r$. Entry via the production of substitutes will be profitable so long as the costs of the entrants are below $P_r$ and above the price which entrants expect to prevail after entry. The costs of the potential entrants are assumed constant at level $AC_e > AC_m$.

Now suppose that the regulated firm is not allowed to retaliate against entry by cutting price in the market segments affected. DD, its demand curve, will shift to the left as entry occurs. The final equilibrium could be with the former monopolist facing demand curve $DD'$. In this case, the best that it can do is earn a rate of return just equal to its cost of capital by charging its monopoly price, $P_m = AC_m$, and producing the corresponding monopoly output $Q_m$ at minimum possible cost. If the regulatory authority values the survival of its ward and is unable to prevent entry, it will allow the firm to move to $P_r$, $Q_r$.

What are the efficiency effects of this episode of entry? Output efficiency is improved since price has fallen from $P_m$ to $P_r$, toward the minimum marginal cost of production, $MC$. The magnitude of this gain is small, given by area $abc$. Cost inefficiency still remains and is, in fact, more severe. The magnitude of extra costs incurred is given by $P_r - AC_o$, $Q^*$, which can clearly exceed the welfare gain $a+b+c$. Of course, if the entrant’s cost, $AC_e$, had been below the initial average costs of the regulated firm, $AC_m$, cost efficiency would have been improved, and the welfare effects of the entry would have been unambiguously positive. In the case portrayed in Figure 4, the final cost inefficiency, relative to a qualified first-best solution in which the regulated monopolist is somehow induced to produce the output $Q^*$, is due to overly small production by the decreasing cost monopolist and production of $Q^* - Q_m$ by the higher cost entrants; the Averch-Johnson bias plays no role here. And, even in the first-best case mentioned here, output inefficiency remains.

This analysis leads to a significant conclusion for regulatory policy. If the concern is to promote economic efficiency, then commissions should be alert to the possibility that inefficient entry, in effect induced by the Averch-Johnson bias, may occur. Unless regulated firms with lower costs are allowed to retaliate, the commission may find itself in the position of having to ratify inefficiency to ensure a steady flow of even normal profits to its clients. Figure 4 illustrated a case in which, on static efficiency grounds, a prohibition of entry
would have been preferable to free entry, but note that this preference only holds on the assumption of a myopic regulatory authority.

**Summary: Entry Restrictions and the Form of Regulation**

The likelihood and effects of entry thus depend not only upon the usual barriers to entry, but also upon the goals the regulators seek and the specific means they pursue to achieve them. If direct entry is possible or if close substitutes are available, regulators should be careful lest their attempts to protect their clients cause increased inefficiency. Regulators should also hesitate before acting to protect firms from themselves by disallowing price cuts designed to meet emerging competition.

**Efficiency, Entry, and the Effects of Regulatory Lag**

**Entry Possibilities and the Choice of the Fair Rate of Return**

The level of the allowed or fair rate of return, \( s \), set by the regulators has complex effects on both input and output efficiency, some of which were discussed in the first section. How do the regulators arrive at a value for \( s \)? A believer in the theory of "regulation for the benefit of the regulated" would hold that the allowed rate set is at a level which allows maximization of profits. A cross-subsidization rationale for regulation might point to the rate being set at a level that could sustain the desired magnitude and extent of subsidy. If regulation attempts to simulate the results of competition, the regulators would presumably fix \( s \) as close as possible to the cost of capital; much of regulatory proceedings is concerned with this effort.

Here the concern is with two related questions. First, at what level should regulators set \( s \)? Second, how might entry possibilities influence their choice? The first question has been addressed at length by Klevorick, in the context of the Averch-Johnson model. His question was: "Should the fair rate of return always be set equal to the cost of capital?" An initial difficulty is that, if regulation is continuous and assures that at each moment \( s = r \), the maximum total profit that the firm can earn is zero, and the choice of inputs and level of output are indeterminate. If the allowed rate of return equals the cost of capital, a goal in addition to simple profit maximization must be added. Goals which put the \( s = r \) solution in perhaps the best light are revenue or physical output maximization, both of these imply cost minimization and input efficiency. To evaluate the desiribility of alternative levels of \( s \), an explicit social welfare function must be used to evaluate the welfare outcome when \( s > r \), the firm choosing input to maximize total output. This outcome can be compared with the welfare level when \( s = r \) and the firm is assumed to maximize profits in the usual constrained manner. Using the total of consumers' plus producers' surplus as a welfare criterion, Klevorick shows that solutions with \( s > r \) could be preferred. In an example using specific forms of the production and demand functions, Klevorick finds that the greater the degree of returns to scale, the closer the optimal fair rate of return moves to the cost of capital and vice versa.

This result can be partially clarified by the graphical illustrations of Figure 5. In each part of the figure, \( AC \) and \( MC \) are the regulated monopolist's average and marginal cost functions, and \( DD \) is the market demand function. If \( s = r \), the firm's costs are just covered and price must be equal to average cost; these average cost prices and the corresponding outputs are given by \( P \) and \( Z \). The market clearing marginal cost price and corresponding output are given by \( P \) and \( Z \). Assume that the regulators cannot allow the firm less than the cost of capital, so \( s > r \). If the welfare criterion is maximization of consumers' plus producers' surplus, the optimal output is in each case \( P \) and the optimal quantity \( Q \).

With constant returns, as in Figure 5a, average and marginal cost coincide. Setting \( s = r \) leads to ideal output and price. In \( 0 \), decreasing returns imply that \( MC \) exceeds \( AC \). Klevorick, in the context of the Averch-Johnson model, was concerned with the industry. For this firm, price should be adjusted upward so that output will move toward \( Z \). This could be achieved by setting \( s > r \).

[13] The model of Figure 4 did not explain why the entrants would not attract the same costs as the initial monopolist. Perhaps the cost disadvantage is a short-run phenomenon which will eventually be overcome by learning. Also, some restriction must be invoked to explain why the monopolist does not solve the entry problem by merging with or otherwise buying out the potential entrants.

[14] Note that attempts to adjust realized rate of return to the cost of capital directly simulate only one competitive result: zero profit, long-run equilibrium.

[15] If \( s > r \), these goals do not result in input efficiency. See Bailey and Malone, pp. 33-36, and Bound and Klevorick, p. 37. The situation of Figure 2 illustrates some delicate issues discussed in 17. The question is whether the cost curves shown refer to the firm or to the industry. If they refer to the firm, then the presence of only one firm in the market must be explained, perhaps most plausibly by entry restrictions.
With increasing returns, as in Figure 5c, $MC < AC$, so $P_r > P_e$, and $Z_r < Z_e$. Achievement of the optimal output would require a realized rate of return below the cost of capital, or some other form of subsidy. Since this is not allowed, the best that the regulators can do without price discrimination is to set $s = r_u$, approaching $Z_e$ as closely as possible. Note that these results are not complete, as the effects of variation in $s$ on input efficiency were not considered. In Figure 5b it is possible that a move to $s > r_u$ could raise costs sufficiently to offset any gain in output efficiency.

The above discussion can illuminate some aspects of regulatory policy toward pricing and entry. In the decreasing returns case, if the cost curves of Figure 5b are for the firm, then its output is too large for efficiency. If smaller firms could be replicated and several could occupy the market, each producing $Z^*$ with price $P^*$, efficiency would be served. Total industry output would then be $Z^{**}$. One way of arranging this might be to set a high allowed rate of return. This could serve rapidly to attract entry of new firms. Why had entry not occurred at $P_2$? One likely reason is regulatory prohibition. Another is that the existing firm might value its degree of market occupancy or level of output and negotiate with its regulators for price decreases if it expected entry. Or, potential entrants may have higher costs. Finally, minimum optimal firm size may be high enough and demand inelastic enough that entry is not considered profitable. If firms cannot be replicated, or for some reason several cannot occupy the market, efficiency would still dictate an approach to $Z_e$ through raising the fair rate of return.\(^{17}\)

In the constant and increasing returns cases, the entry of additional firms would, within the confines of the Averch-Johnson model, serve no efficiency function and thus should not be encouraged by allowing returns in excess of the cost of capital. Of course, entry might be valued because of its possible beneficial effects, such as on innovation and internal efficiency of firms.

Examination of the likely behavior of the regulators complicates the situation. Regulators may prefer to allow their constituents excess returns because they are captives of the firms, because they want to maintain cross-subsidization, or because they prefer to hear less complaining from their close associates. However, regulators presumably do not like to be besieged by applicants for new franchises.

\(^{17}\)Perhaps firms could be replicated only at much higher cost. One might also conceive of external effects such that with several firms in the market, the costs of all would be considerably higher.
Their desire to allow excess profits will then be directly related to the difficulty of entry. And, especially if entry can occur via increases in the output of unregulated firms, the ability of regulators to allow excess returns will be similarly affected. A paradox then exists. In situations of increasing returns, full efficiency would dictate a price below average cost. But with significant scale economies, entry is likely to be difficult and the regulators willing and able to allow excess profits. By contrast, in situations of decreasing returns, a high allowed rate of return might be beneficial. Yet, entry is likely to be easy, and the regulators less likely and less able to allow substantial excess profits. 18

6. Suppose the firm initially produces at A, assumed to be input efficient. Profits are lower than at all points on S, the constraint curve, that lie between A and D. Looking ahead to a regulatory review and price adjustment at time T, the firm reduces labor by an amount $\Delta X_l$, raises capital by $\Delta X_k$, and moves to point C. Its profits are temporarily lowered. At time T, prices are adjusted so that the realized rate of return equals $r$, and the firm moves to point F on the constraint curve, enjoying higher total profits than at A.

![Figure 6. Effects of Regulatory Lag](image_url)

The implicit assumption of the model is that from time T onward, the firm remains at F. On this assumption, Bailey and Coleman show that for short regulatory lag periods the firm would immediately overcapitalize by the full Averch-Johnson amount, moving im-

---

18Presumably, the regulators can effectively control direct entry into the regulated areas. The ease of effective entry will then depend upon the economic barriers to entry into regulated industries which produce close substitutes.

19See Bailey and Malone.
mediate to point D. For longer lags the amount of overcapitalization will be less than the full Averch-Johnson amount. Yet, suppose that the firm is at F. Looking ahead to another regulatory review, the firm would again increase its capital and decrease labor, temporarily lowering its profits. After review, it would end up with higher total profits, at some point closer to D. Now, suppose that the firm is at D. Output is temporarily fixed at E. If E is the cost minimizing point for this output, the firm can benefit, at least in the short run, by moving toward E, the regulatory authority will then lower price, increasing output and moving the firm back toward output $Z_c$. There seems to be no reason in this model why the process might not continue. Costs of adjustment of price and input mix could certainly result in an equilibrium being attained. In many utility sectors capital is extremely long-lived and immobile, so input proportions cannot be adjusted cheaply. While the final equilibrium point cannot be specified, the average degree of input efficiency and level of output will exceed that at D, the equilibrium point with continuous regulation.  

The extent of input and output efficiency depend on entry possibilities. With lagged regulation, firms overcapitalize in the expectation of larger future allowed and realized profits. Once price is adjusted upward, entry could easily reduce or eliminate the hope for gains. So, if firms correctly forecast the effect of increased prices and profits on entry, there will be a reduced incentive, or overcapitalization. If, as discussed above, entry serves to keep the maximum profit that the regulators can allow close to the cost of capital, the regulated firm may have no choice but to be efficient. In this case the length of regulatory lag would be of less importance.

**Regulatory Lag and Innovation**

If the allowed rate of return is held close to the cost of capital, innovation in a context of regulatory lag might be the major source of profits. At the same time, regulatory restrictions on entry guarantee that at least some of the benefits from innovation accrue to the regulated firm. The situation is depicted in Figure 7. The CC function measures the present value of the cost of developing an innovation as a function of development time. BB measures the present value of the firm’s returns from the innovation. The downward slope of the benefit function reflects the fact of discounting and also the probability that, as time passes, it becomes increasingly likely that a rival will develop the innovation.  

![Figure 7: Benefits and Costs of Varying Speeds of Development](image)

An increase in regulatory lag will have two effects. It will shift BB as well as change its slope. First, BB shifts out, as the firm can expect to enjoy the profit benefits on an innovation for a longer time. This will increase the likelihood that any given innovation will meet the requirement that total benefits at least equal total costs, so more innovations will be made, increasing the rate of technical progress.

---

80This might not be the case if regulatory lag allowed the firm to charge its desired monopoly price or if shortages developed.

81This model is based on one presented by Scherer, pp. 396 - 70.
At the same time, the increase in lag means that consumers will receive a lower proportion of the benefits, and stockholders a higher proportion. Entry prohibitions are likely to affect both the position and slope of BB. An entry prohibition will shift BB outward because there will be fewer rivals to divide the possible total benefits from an innovation.

At the same time, however, the slope of BB will be lower than with unrestricted entry: there will be less risk that delaying development will result in a real making the innovation. This effect could lengthen the typical optimal development time and slow the rate of technical progress.

In fields where the opportunities for profitable innovation are small, in the sense of high CC relative to BB, entry prohibitions may serve to accelerate technical progress by assuring a favorable cost-benefit ratio for a larger number of innovations. In a regulated industry with $s - r$, this stimulating effect might only be observed if regulatory lag were sufficiently long. In fields where the technical opportunities are high, the effect of entry restrictions could be to lower the rate of innovation by flattening BB. This would lengthen the optimal development time for innovations which might have met the profitability test even in a situation of unrestricted entry.

Determinants of Regulatory Lag

While regulators occasionally review some cost items and generally review proposals for new construction, the discussion so far portrays them as using two basic means of control: the allowed rate of return, $s$, and length of regulatory lag, $T$.[25] As discussed above, the allowed rate of return will depend both upon what the commission wishes to give and upon the returns it is able to guarantee. The period of regulatory lag similarly depends upon the goals of the regulators and pressures and constraints they face.

For given regulatory resources the frequency of reconsideration of the prices charged will vary directly with the number of demands for review by the regulated firms. In sectors where increasing returns are the rule, barriers to entry will tend to be high, and the regulators will be able to set and sustain a high allowed rate of return. Furthermore, there are likely to be fewer firms, fewer petitions to enter, and fewer disagreements among firms as to the most appropriate prices. Regulatory review of pricing would then be relatively infrequent, and the average regulatory lag long.[25] While the high allowed rate of return is likely to contribute to output inefficiency, the long regulatory lag will help alleviate any tendencies to cost inefficiency.[25]

On the contrary, in decreasing return areas, the number of firms is likely to be large and entry relatively easy. The regulators will be less able to set a high allowed rate of return, and firms are more likely to be dissatisfied. With more firms, pricing disagreements and jurisdictional disputes over market areas will be more likely, and thus pricing reviews more frequent. The average lag should then be relatively small, and the incentive to overcapitalize greater. Yet, this tendency will be somewhat constrained by the fact of easier entry.

Conclusions

The effects of entry controls and their relaxation depend upon the specific methods of regulation and the market situation in question. A variety of arguments indicate that, under some conditions, relaxation of entry controls might worsen both the input and output efficiency of regulated firms. In conditions of low opportunity for profitable innovation and small economic barriers to entry, an open entry policy could slow the rate of technical progress.

From the standpoint of efficiency, unrestricted entry into regulated sectors seems least ambiguous in effect when the likelihood of an Averch-Johnson cost bias is low. Firms might lack the detailed knowledge of technology and demand needed to implement the bias fully and profitably. Production technology might allow so little input substitution that any cost is insignificant.

If relaxation of entry controls is to increase efficiency, regulators should generally allow their wards to respond to the pricing initiatives of entrants. If entrants' costs are likely to fall as they become more familiar with the technical, market, and regulatory environment, some temporary "infant industry" protection from price cut-

[25] Review of new construction might be seen as an attempt to prevent "year-lose padding." Peters (p. 35-36) argues that these controls are best interpreted within the framework of a cross-subsidization theory of regulation.

On the contrary, one might argue that the feebleness of firms would imply that the authority could afford to conduct more frequent reviews. The question could only be resolved by some explanation of the size of budgets and the operating efficiency of regulatory commissions.

At least the lag would alleviate the Averch-Johnson bias. The absence of regulatory scrutiny combined with relative freedom from entry threats could breed other inefficiencies.
ting by established firms could be desirable. This provision should not be permanent.

Even if large Averch-Johnson cost biases were likely, and even if open entry moved output in an inefficient direction, unrestricted entry might still be preferred. In fields of high technical opportunity the policy is likely to result in more incentive to innovate. Even if the market structure remained highly concentrated, the actual or potential presence of rivals would constrain cost inefficiency.

In some sectors an open entry policy might result in a significantly more atomistic market structure, perhaps weakening the power of any one firm vis-à-vis its regulators and also lessening the need for regulation. Yet, if firms become more numerous, the regulatory task could be rendered so much more difficult that regulation would become not less restrictive, but more arbitrary, rigid, and prone to entry.

References


Michael Klass
III. General Conditions of Entry

William G. Shepherd

Under pure competition, all relevant competition is actual. At the opposite extreme, in pure monopoly, the only relevant "competition" is potential. For regulated utilities in particular, potential competition may be important. Yet, the regulated firm not only is a monopoly, but also has a monopoly franchise, by which the state itself guarantees to exclude potential competition. Monopoly under regulation is doubly complete, and potential competition may be doubly important.

As discontent with commission regulation of utility firms has grown in recent years, these formal entry barriers under regulation have come under special criticism. Open up entry, it is said, and monopoly utilities will tend toward competitive behavior in pricing, efficiency, and innovation. Regulation could then recede to its more basic functions, shedding those attempts to control profit and price which probably create more costs than benefits.¹

This strategy will be efficient if a lowering of barriers does actually lead to changes in structure and behavior in the "utility" market.

General Conditions of Entry

Yet, that view assumes that the regulatory barriers are the only ones in force and that they do influence behavior. If either premise is false, the man-made regulatory barriers might be removed, but without leading to any increase in competition. Instead, strictly economic barriers might remain quite high enough to exclude new competition. Or it may be that entry barriers are simply not an important influence on market activity at all, in utility sectors and elsewhere.

One therefore must begin on the "entry question" at the very beginning, with fundamental inquiries about what entry and entry barriers really are. This will be done here. It will be seen that there are two main kinds of entry to be analyzed: entry of new sellers and of new management. The next three sections will outline, for purposes of comparison with utility markets, the actual effects of entry barriers in industrial markets. Unless entry is a powerful force in the broad cross-section of industrial types, it hardly can be expected to play a strong role when one tries to introduce it into regulated sectors. Econometric tests will show the average relationships which hold in intermediate ranges of industrial market structure, and case studies will help to explore the border conditions of large-scale entry, complete nonentry, and entry by take-over. This industrial evidence suggests the general conditions which maximize the net benefits from actual and potential entry. These, in turn, suggest strategies which might be pursued in fostering entry into regulated utility markets.

Concepts of Entry and Barriers

There are two basic forms of entry. One is the insertion of a new competitor among the existing sellers. This, to economists, is the conventional concept of entry: Wilkinson Sword, Ltd., enters the razor blade market, IBM enters copying equipment, and so forth. The other basic form is new competition for the function of managing and existing seller. This occurs commonly by a take-over, or related strategy, and it most often is directed at firms whose managers are performing inefficiently, by any one of several criteria. The conceptual distinction between entry as seller and entry as manager is important, even though the two kinds often are combined in practice. The two types have different causes, form, and effects, and often one is feasible even though the other is not. This fact can be important for entry into utility sectors.

The literature has dealt in some detail with the first form of entry. Potential entry was advanced as a critical market factor by J. B. Clark as long as 70 years ago. It was asserted to be capable of neutralizing all market power, even that of a complete monopolist. Although it was not defined or analyzed precisely, potential entry presumably meant that the monopolist’s demand curve was highly elastic at prices above the level which yields normal profit. That is, an attempt at monopoly pricing along an inelastic demand curve would convert potential into actual competition and switch the former monopolist onto an elastic portion well to the left, as shown in Figure 1.

![Figure 1](image-url)

**Figure 1.**

- Former "monopolist" ex post demand curve, after potential competition becomes actual
- Monopolist ex ante demand curve
- Marginal cost
- Average cost
- Marginal revenue
- Quantity

---

1. The most thorough treatment of entry and barriers is in Joe S. Bain, Barriers Against New Competition (Cambridge, Mass., Harvard University Press, 1956), chapter 1. Bain deals with "Type I" entry, as defined here.


supplants the variable for market share, the extreme barrier hypothesis is supported (that is, barriers are the critical element of market structure).

Yet, it is entirely likely that such substantial entry is only a border case, normally of little concern. One first must define it precisely; Bain maintains that substantial entry involves a net addition to industry capacity by a new firm previously not in the industry. This firm’s addition must exceed the sum of all other firms, which its entry may cause. If that exit is by many small, faceless, fringe firms, while the entrant is a single, large entity, the whole share of the entrant may be regarded as “net” entry by the leading firms. But that point leads one to redefine entry as entry into the leading firm group rather than into the whole industry, and such “entry” into the leading group is, strictly, turnover rather than market entry. This type of “entry” also could occur by rapid growth of established small firms (possible branches of large diversified firms), or by strategic intrusions by middle-range firms. In fact, the impact on each leading firm from strategies by other leading firms (for example, new product innovations and advertising campaigns) would be essentially the same as from such “entry.”

Properly defined, entry into an industry is likely to have at least three main dimensions: (1) extent, for example, the market share which is acquired or is believed to be the entrant’s target, minus the actual or expected exit by identifiable competitors; (2) speed of occurrence; and (3) distance, or degree, of surprise. Only a case which has high values for these attributes can be regarded as significant entry. Lesser values are indistinguishable from the host of little moves and random shifts by already established firms. These small disturbances in turn shade into the anonymous background of competitive constraints on the firm.

Yet, the possibility of occurrence is inversely related to precisely
these values.\textsuperscript{11} Indeed, firms commonly adopt strategies toward precisely the probability of entry. And suddenly, net entry, rapidly and on a large scale, into a firm's market area is, by definition, improbable, as are the wide variety of other possible calamities. To the extent that entry, or of any change, is the short-run reduction of the firm's market share (and of future profit expectations). In this context, "limit pricing" relates to the share-maintaining price strategy of the individual firm, rather than to the joint pricing strategy of leading firms towards possible entry into "the" industry.

Entry into the industry, therefore, tends to be indistinguishable in concept from all of the other competitive constraints on the choice set of the firm, and large-scale instances of it are normally of second-order likelihood.\textsuperscript{12} As for the "conditions of entry" or "barriers" which researchers such as Bain, H. M. Mann, and others have tried to estimate, they have tended to emerge as amalgams of specific influences which are not easy to formulate or measure. Bain has argued that the three main sources of entry barriers are probably scale economies, specific absolute cost advantages, and product differentiation. Each is relevant to certain industries (or firms within them), but each is important only to relatively few, and these are generally not overlapping sets. We have, therefore, a series of different situations, not so much a general phenomenon. Other ad hoc barriers also have been suggested, such as patents and research and development, and still others could be suggested. Attempts to weld such instances into objective measures of barriers have yielded interesting but rather vague and untestable classifications.\textsuperscript{13} The alternative is to focus measurement and analysis on the possible "barrier" factors separately and precisely (as with size and advertising-to-sales ratio) in analyzing the firms for which they are likely to be influential.

I now turn, more briefly, to entry by take-over, which introduces new competition for the management function within the existing firms. Take-overs range from fine degrees of partial take-over (such as placing representatives on the board) to complete control and displacement of the previous management. The conditions under which mergers may arise are complex, of course, and take-overs are a border case of nonamicable merger. The determinants of take-overs also are not simple,\textsuperscript{14} but there are several common features. First, the present management usually is receiving and acting on its opportunities at levels below reasonable standards of performance; that is, it is inefficient. In few cases are normal or highly efficient managers a take-over target. Second, this gap in performance is perceived by an outside group which possesses management resources capable of assuming new responsibility. Third, this group secures adequate financial support, either from specific banks and others or, if it is already a large enough firm or holding company, from its own resources. These resources may need to be very large. Fourth, a tender offer is made, usually unexpectedly and at about 20 percent above the current market price of the stock (proximity of the offer is less frequent and less likely to be successful).\textsuperscript{15} Fifth, there is a large variance in the degree of success of new management, but it is commonly above the preceding level. Sixth, the trauma of a threatened or nearly successful take-over usually is severe and results in substantial improvement in "old" management's performance, frequently accompanied by major revisions in personnel. In short, take-overs — actual and threatened, direct and indirect — commonly raise the level of managerial performance.
The Role of Barriers in Large-Scale Industry

The next three sections offer evaluations of the role of barriers and seller entry into a range of manufacturing industries in the United States. Such industrial patterns cannot, of course, show directly what would happen if seller entry into utility markets were opened up. Regulated utility industries may be different, either inherently or because of regulation. Also, what has happened in the past may not correctly show what would happen in the future.

Moreover, we might reasonably expect that entry into regulated utility sectors will be less rapid and effective than it is into other markets, such as manufacturing industries, which are more "normal" in technology and demand. Therefore, industrial evidence that barriers are influential — that is what we find — would not establish that the same would hold for regulated sectors. But contrary findings—that barriers are not important in industry — would give a relatively stronger indication that they would not matter very much in regulated industries either.

The other reservation about the regressions in this section is that the data cover firms with market shares below 75 percent. There simply are very few firms with shares above 60 percent in industrial markets. If utility firms have market shares genuinely above that, and many if not all of them do, then we cannot be sure that the industrial results are fully (or even partly) transferable. This may be true of some utilities but not others.

Nonetheless, industrial patterns are the most general basis for predicting entry behavior in the economy. They reflect the role which banks and other sources of capital play toward new entrants in any sector, and they represent a diverse and substantial share of the economy. As noted, the firms discussed here are large ones, at least comparable in scale to those in most regulated industries. These industrial patterns are the only real source of general information which we presently have, a minimum but perhaps persuasive merit.

William Shepherd

Preliminaries aside, what are the industrial patterns? Barriers could be important in two main directions: They could determine profitability, and they could govern the vulnerability of established firms\' market shares to erosion over time. Both of these possibilities will be examined in turn.

Profitability

There have been statistical hints from Bain and Mann that barriers do influence profit rates. These hints derived from average profit rates tabulated for leading firm groups in 30 industries. The industries were divided into three categories of high, medium, and low entry barriers. The degree of concentration in the industries also seemed to influence profit rates jointly with barriers, but this influence was not separated into a partial association distinct from the role of barriers. Nor, probably, could it be factored out with the data then used.

It is now possible to test the patterns more precisely, using data developed for a panel of 245 of the largest U.S. industrial corporations for which market shares can be estimated reliably. These firms

8Bain, Barriers, and Mann, "Seller Concentration."

The population is the largest 500 U.S. industrial corporations, whose size and profitability are available on a relatively standardized basis in the annual Fortune Directory of the 500 largest U.S. industrial corporations. Their aggregate importance in the U.S. economy is evidently large. The years 1960–1969 provide a recent, relatively recession-free period which is long enough to show equilibrium patterns for many firms.

Coverage. The data cover primarily those firms among the largest 500 manufacturing corporations whose average market shares in the early 1960s can be estimated and which were not involved in a major internal or industry-wide disequilibrium. An additional 14 from the next largest 500 were also included for 1960–1969, making a total panel of 245 firms.

The panel includes a wide range of industries and of market shares — from high to low — within them. Firms were excluded mainly for (1) a high degree of internal diversification, (2) a major merger during the period, (3) high sales to the military (which would particularly affect post-1964 results), (4) absence from the largest 500 group for more than one year during 1960–1969, and (5) major disequilibrium in the firm's primary industry or in its own condition.

The premise for the sample selection is that all firms and subsidiaries conform to the general structural patterns so that these cases with identifiable market shares are representative of the universe. Diversified firms simply render unobservable the individual results of their subsidiaries, which are in other respects as if drawn from the same population as the 245-firm panel. This assumption, if it is valid, mitigates

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Constant</th>
<th>Market share</th>
<th>Share of the other 3 leading firms</th>
<th>Growth</th>
<th>High barriers</th>
<th>Medium barriers</th>
<th>R²</th>
<th>Number of firms</th>
</tr>
</thead>
<tbody>
<tr>
<td>All firms Rate of return on equity, 1960 – 1969</td>
<td>4.861</td>
<td>.251</td>
<td>.001</td>
<td>.921</td>
<td>2.453</td>
<td>1.549</td>
<td>.528</td>
<td>245</td>
</tr>
<tr>
<td>(4.60)</td>
<td>(11.81)</td>
<td>(.05)</td>
<td>(4.36)</td>
<td>(2.84)</td>
<td>(2.64)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All firms except those in industries with high barriers Rate of return on equity, 1960 – 1969</td>
<td>4.756</td>
<td>.239</td>
<td>-.001</td>
<td>.971</td>
<td>1.688</td>
<td>.518</td>
<td>202</td>
<td></td>
</tr>
<tr>
<td>(5.14)</td>
<td>(10.85)</td>
<td>(.03)</td>
<td>(4.29)</td>
<td>(2.03)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All firms except those in industries with high or medium barriers Rate of return on equity, 1960 – 1969</td>
<td>5.392</td>
<td>.375</td>
<td>.019</td>
<td>.725</td>
<td>.461</td>
<td>71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4.23)</td>
<td>(6.14)</td>
<td>(.08)</td>
<td>(3.33)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: t ratios are given in parentheses. All regressions are weighted.

Table 2. Basic Analysis of Profitability of Large U.S. Industrial Firms, 1960 – 1969

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Constant term</th>
<th>Market share</th>
<th>Concentration</th>
<th>Group</th>
<th>Size (log. of assets)</th>
<th>Advertising intensity</th>
<th>Growth</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Profits rate 1960 – 1969</td>
<td>5.13</td>
<td>.250</td>
<td>.028</td>
<td>.182</td>
<td>.051</td>
<td>.792</td>
<td>.554</td>
<td></td>
</tr>
<tr>
<td>(5.62)</td>
<td>(12.19)</td>
<td>(1.60)</td>
<td>(1.34)</td>
<td>(4.39)</td>
<td>(3.93)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Profits rate 1960 – 1969</td>
<td>4.38</td>
<td>.105</td>
<td>.175</td>
<td>.304</td>
<td>1.28</td>
<td>.344</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4.22)</td>
<td>(5.57)</td>
<td>(1.06)</td>
<td>(5.04)</td>
<td>(5.31)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: t ratios are in parentheses.


<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in market share</td>
<td>.566</td>
<td>.44</td>
<td>.44</td>
<td>.962</td>
<td>.039</td>
<td>.936</td>
<td>.54</td>
<td>.120</td>
<td>.79</td>
</tr>
<tr>
<td>(1.53)</td>
<td>(1.55)</td>
<td></td>
<td>(4.31)</td>
<td>(4.21)</td>
<td>(4.31)</td>
<td>(1.12)</td>
<td>(4.31)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change in market share</td>
<td>2.622</td>
<td>.347</td>
<td>.233</td>
<td>.002</td>
<td>.019</td>
<td>.961</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(.777)</td>
<td>(.43)</td>
<td>(.49)</td>
<td>(.02)</td>
<td>(1.65)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change in market share</td>
<td>-2.864</td>
<td>.168</td>
<td>.523</td>
<td>.013</td>
<td>.125</td>
<td>-.120</td>
<td>- .006</td>
<td>.79</td>
<td>.093</td>
</tr>
<tr>
<td>(1.49)</td>
<td>(.21)</td>
<td>(.97)</td>
<td>(6.45)</td>
<td>(5.43)</td>
<td>(5.43)</td>
<td>(1.12)</td>
<td>(4.42)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: t ratios are in parentheses.
group. Advertising intensity clearly is important, but of course it matters only in the seven or eight industries where advertising-to-sales ratios are high. The negative coefficient for size is puzzling. Although it is statistically definite, it implies only a mild reduction of profit rate as size increases. Pending further analysis, it seems most likely that it reflects some degree of inefficiencies of bigness which just offset any positive entry barriers effects on profit rates.

**Market Shares**

Do entry barriers shield high market shares? Conversely, where barriers are low, do high market shares tend to dwindle more than they do in high barrier markets? The industrial panel of 245 firms again provides a relatively clear and broadly based set of data for testing these possibilities for the 1960-1969 period. Some 88 of the 245 firms appear to have had significant changes in market shares during that time. The analysis tests whether the height of entry barriers is related to these changes, both in simple regressions and in multiple regressions including other possible influences on market shares. As before, the entry barriers can only be represented by dummy variables, so that the answers are by no means final.

The results are shown in Table 3. A simple regression (number 1) shows no perceptible association whatever between barriers and market share changes. In regression 2, initial year market share is added, since there are strong indications from other analyses that initially extreme shares (high or low) do regress toward the mean over time. That is, unusually high market shares exhibit a "natural" tendency to dwindle. Thats tendency emerges here, too, but it leaves unchanged the apparent irrelevance of entry barriers.

A fuller analysis is needed, since at least two more measurable influences may be present. One is initial year profitability; when it is high, it may attract more intense competitive efforts by other firms, causing the firm's own market share to recede. The other is profitability during the rest of the period, which has been shown to have a strong positive association with changes in market shares. This association, in fact, strongly suggests that firms strive to maintain market shares, rather than yield them up voluntarily over time as part of a long-run profit maximizing strategy.

In this expanded analysis (the third line in Table 3), barriers are still not associated with market share changes. There are slight negative coefficients where positive ones would be expected, but they are not close to being significant either statistically or substantively.

On this evidence, one can conclude provisionally that barriers are not related to the stability of market shares. Such a null conclusion can never be definitive, since we may simply have failed to measure either variable accurately, but it is the most reliable scientific indication we now have, judging both by the breadth of the interindustry coverage and the completeness of the model. Expressed another way: There would have to be startling changes in the underlying data in order to alter the results enough to suggest a powerful influence of barriers upon market share stability. Pending that, it seems reasonable to presume that high and moderate barriers do not affect the vulnerability of the market shares of established firms across the range of large-scale industries in the United States.

The strong partial association between market share and profitability also indicates that changes in market share are a powerful determinant of company success. To the extent that it reduces market shares, entry will exert a strong influence on established firms' behavior, but only to that extent. Trivial, slow, or peripheral entry will scarcely affect established market shares and behavior.

Translated into regulatory policy, this means that entry must be large. An opening of entry must be so substantial that remaining barriers are negligible, and it must sharply depress market shares in the utility firms' prime markets in order to apply an effective constraint. Such entry would be on the order of at least 15 to 25 percent, and it would be rapid. The utility's interest will lie in portraying entry, no matter how tiny, as threatening to have a great impact on its market share and its viability. But regulators should be guided instead by the objective indications, which are that entry must reduce market shares sharply below 100 percent — and in the long run probably below 50 percent — in order to provide a real substitute for regulatory constraint.

Moreover, the share reduction must be permanent. Only at market shares below 40 percent does the rate of return approach competitive levels (Tables 1 and 2). A share which drops temporarily to, say, 75 percent and then rises toward 100 percent will continue to yield profits at least as high as 30 percent. This is a vector of two influences. The 75 percent share would itself normally yield rates of return above 20 percent, and the additional effect of rising market share is strongly positive.

There are other cross-section data suggesting that entry alone is not likely to be critically important. Charles Berry reports that most entry in U.S. industries in the 1960s was of a creeping sort, with low
values for the three dimensions of entry: size, speed, and surprise. This is confirmed by the survey of major cases of entry since 1955 in the next section. There were few cases of drastic entry: All required at least two years to occur, none involved genuine surprise, and some arose from deliberate public policy actions (as in aluminum and synthetic rubber). Most "entry" in fact occurred by sharp rises in small established firms.

Japanese concentration data, which are extremely detailed, offer another test. The top 15 firms are listed in successive years, which should make it possible to identify specific entrants into industries over those years. But an attempt to do so by this author for 169 industries failed; it proved impossible to distinguish such cases of "new" entry from existing small firms and turnover among all significant firms.

In short, the normal experience of industry is a variety of small changes with small or intermediate effects. Among these, new entry mingles with other shifts. If entry barriers influence profits, they operate via market shares; a direct barriers-profitability relationship has not yet been observable. Only if the reduction of barriers and the subsequent market entry are substantial and lasting — toward market shares below 50 percent — will they constrain utilities to approximate standards of competitive behavior.

Cases of Seller Entry

Entry is a distinctive event; it is not small firm growth, intra-firm oligopoly shifts, or product innovation. It is not new capacity added by a new firm. An extensive search of industrial sources was made in order to identify the main cases of entry in U.S. industries since 1955. Remarkably, only a handful of clear, significant entry cases emerged, despite common rhetoric about dynamic behavior among large corporations. As Berry shows, most "entry" is so incremental as not to be entry: clear, rapid, substantial entry by new firms. In the

light of earlier discussion, this should not be surprising. But the apparent paucity of known cases of major entry since 1955 is more striking than one would have expected.

Presently, a definitive count is impossible because the key data are unavailable, thanks to census disclosure rules. The known prominent cases include: Armour & Co. into hand soaps (after 1958); Wilkinson Sword, Ltd., into razor blades (1962); Procter & Gamble into paper (1962–1968); Iowa Beef Packers into meat packing (1961–1965); and IBM into copying equipment (1971). In addition, imports have risen in several industries during the 1960s, notably steel and automobiles; this, too, is a form of entry.

Note that the entrants as a rule are (1) already established firms or (2) imports. The pure newcomer does occur occasionally; Iowa Beef Packers being one, but for major entry one looks usually to large, existing firms. Note also that entry can cause a sharp drop in profit rates and stock prices, but that this does not always occur (exception: Xerox, General Motors).

Several lessons about entry and its probable effects can be drawn. First, markets usually experience a series of small entries, exits, and shifts among fringe members, as well as moderate adjustments among the larger firms. Even the large cases of "entry" are often indistinguishable from the flow of small changes during a long initial period. Many cases of supposed "entry" are merely rapid rises of smaller, new established firms, such as Pepsi versus Coca-Cola in the 1950s.

Second, firms with high market shares frequently do become sluggish, thereby inviting such kinds of entry. However, the leading firms in the very largest markets (oil, automobiles, steel, computers, electrical goods) appear to have become virtually exempt from entry by take-overs. Third, dominant firms commonly do not respond to entry until it is large and rapid, in the range of 15 to 25 percent of the market. The trigger is the perceived loss of market share. Often the entry has proceeded for many years before it is adequately perceived and responded to by the prior firm (for example, Scott paper and Procter & Gamble; steel firms and imports).

Fourth, entry on a large scale (15 to 25 percent of the market)

50 General Conditions of Entry

William Shepherd 51
Six of these 13 industries appear to have had no important new entry during 1964–1967, judging by a survey of industrial, trade, and financial sources. A seventh, metal cans, reflected the workings of an important anti-trust consent decree and therefore was made. Moreover, the entry was primarily done vertically, by established can users. In the remaining six industries, significant new entry probably occurred. These accounted for about 5 percent of the total shipments in the panel of 90 “large” industries. But even in these, entry was not always a clear cause of the drop in concentration. Therefore, new entry cannot be said to have been a strong depressant on concentration in major industries. Rather, one can say that in a few cases it is associated with a dwindling of the leading firms’ market share. In other cases, entry occurs without clear effect. In still other cases, entry does not occur even though inducements for it are very strong.

**Absence of Entry into High-Profit Industries**

Several prominent markets in which entry did not occur despite the inducements of high profitability are summarized in Table 5. Each involves high, sustained profitability by the existing firms and a virtual absence of entry, or in some cases an excess of exit over entry. Again, the coverage is neither exhaustive nor a random sample; the research base for either approach simply is not available. Instead, these cases are used to explore monopoly into major industries in which details are fairly fully known.

There are three main lessons to be gleaned from Table 5. First, effective barriers against entry are usually specific devices or strategies, not general conditions. To a large degree, barriers are man-made and artificial. The devices include patents, copyrights, and sometimes legal barriers to entry (as with Western Electric), and such utility buying practices as “Buy American” policies. The strategies include price discrimination by IBM and threatened retaliation against the sole supplier (Baykar Kodak and Polaroid).

Second, entrants fail where the existing monopolist has a network of technology or a sales force which sustains its large market share. The market share itself is vulnerable without such a basis. In particular industries and heterogeneous catch-all “industries” were excluded as being nontypical. Where data were not available for all of the 1964–1967 period, the 1965-1967 trend was extrapolated backward in line with information from a variety of other sources.
lar, the existing firm's ability to price discriminate at or below short-run marginal cost against entrants is often sufficient to deter entry.  

<table>
<thead>
<tr>
<th>Industry and leading firm</th>
<th>Approximately percentage rate of return on equity 1960 - 1969</th>
<th>Device limiting entry</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Telephone equipment</td>
<td>10.0</td>
<td>Roll system purchasing policies and 1966 antitrust consent decree</td>
</tr>
<tr>
<td>Western Electric</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Computers</td>
<td>17.8</td>
<td>Price discrimination</td>
</tr>
<tr>
<td>IBM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Photographic supplies</td>
<td>20.1</td>
<td>Patent</td>
</tr>
<tr>
<td>Eastman Kodak</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polaroid</td>
<td>19.1</td>
<td></td>
</tr>
<tr>
<td>4. Heavy electrical</td>
<td>15.0</td>
<td>Utility purchasing power</td>
</tr>
<tr>
<td>equipment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>General Electric</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Canned soup</td>
<td>12.5</td>
<td>Shelf space limiting price?</td>
</tr>
<tr>
<td>Campbell Soup</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


Third, reciprocal entry by the existing firm against an entrant with a primary basis in another market can often punish and stifle the original entry. In major cases of attempted entry, such retaliation (actual or threatened) is common. Only when the entrant is substantially larger than the prior firm in the entered market will such retaliation fail to deter entry; IBM as an entrant against Xerox illustrates this point.

**Entry into the Management Function**

Information on take-overs is rich but difficult to codify and analyze. Here the experience of two firms which have engaged extensively in take-over activity since 1960 will be drawn upon: ITT and LTV. Several take-over attempts which failed are instructive, as well as others which succeeded. The distinction between take-over and amicable merger is often unclear, since target managements frequently will resist only in order to raise the offer price.

Several aspects of the successful take-over of leading firms — which is the closest parallel to the take-over of utilities — are best illustrated by ITT and LTV. As Table 6 shows, ITT usually has selected leading firms in recent years and LTV has done so as well. As a rule, the target firms had internal inefficiency; thus, ITT expected to be able to reduce the costs of purchased goods by 20 percent purely by harder bargaining on input prices, an addition to other improvements. Both firms did, in fact, improve management performance in many cases without large-scale managerial removals. LTV kept many of its subsidiaries separate financially, with full disclosure of information. The lesson here is that there are large enterprises capable of perceiving, and of acting to correct, the deficiencies in performance of firms with large market shares. These improvements do not necessarily require drastic changes in the target firm's personnel.

Some target firms were as large as most major utility firms, whether one measures size by employment, sales, or aggregate stock value. Therefore, the potential take-over of virtually all the eligible utility firms is well within the scope of past experience in the industrial sector.

Take-over attempts frequently failed during the last decade. Among the many conditions for such failures, two were common. The

<table>
<thead>
<tr>
<th>Firms</th>
<th>Main industry</th>
<th>Approximate position</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITT</td>
<td>Housing, Callous, and so forth</td>
<td>Largest</td>
</tr>
<tr>
<td>Levitt &amp; Sons</td>
<td>Vehicle rental</td>
<td>Largest</td>
</tr>
<tr>
<td>Rayonier</td>
<td>Airport parking</td>
<td>Second largest</td>
</tr>
<tr>
<td>Avis</td>
<td>Sheraton Hotels</td>
<td>Largest</td>
</tr>
<tr>
<td>Airport Parking</td>
<td>Hotel</td>
<td>Second largest</td>
</tr>
<tr>
<td>Continental Baking</td>
<td>Insurance</td>
<td>Largest</td>
</tr>
<tr>
<td>Hartford Fire Insurance</td>
<td></td>
<td>Leading</td>
</tr>
<tr>
<td>LTV</td>
<td>Must, sporting goods, drugs, holding company</td>
<td>Among the four largest</td>
</tr>
<tr>
<td>Wilson &amp; Co.</td>
<td>Boeing</td>
<td>Braniff is largest on its air routes</td>
</tr>
<tr>
<td>Great Northern</td>
<td>Jones &amp; Laughlin</td>
<td>Sixth largest</td>
</tr>
<tr>
<td>(includes Braniff Airways)</td>
<td>Steel</td>
<td></td>
</tr>
</tbody>
</table>

first was when the aggressor firm was smaller than its target and of 
more recent vintage (LTV's failure with Jones & Laughlin is an example). The target firm could then successfully portray its oppo-
nent as an unreliable manipulator. If this persuaded the aggressor’s 
financial backers, the take-over attempt immediately collapsed. The 
second condition was access to a public agency or policy as a way to 
shield the target firm. Thus, Armour in 1968–1971 invoked the 1929 
Meatpackers Decree as a device to persuade the Department of Jus-
tice to seek to prevent General Host, and later Greyhound, from 
acquiring it. In 1968–1969 Goodrich Rubber Co. resorted to a variety 
of quasirelevant legal devices to stave off Northwest Industries. 
The lessons to be learned here are that public agencies (including 
regulatory commissions) are easily used to prevent take-over; neu-
trality will usually require deliberate efforts by the agency to prevent 
such use. Also, regulators need not expect a rash of merger attempts 
by “irresponsible” firms. These would, in any event, commonly fail to 
gain adequate financial backing or would otherwise be averted by the 
target firms. In short, past experience suggests that a supply of 
acceptable take-over agents for laggard utilities would be forthcom-
ing if regulatory authorities preserved strict neutrality.

Conditions Affecting the Net Benefits from Entry

One can now summarize the main lessons about entry. For the 
first type of entry — as a new seller — the lessons are:

1. The effects on prior firms occur via the reduction of market share 
and the rise in uncertainty about future behavior.

2. Entry is easier if the technology is open and if entrants have 
access to data and to supplies of equipment. Vertical integration 
inhibits both of these conditions.

3. Entry is easier if the established firm’s markets are homogene-
ous, rather than varied between profitable core markets and 
unprofitable fringe activities.

4. Entry has market effects only as the “entrant” rises above 15–25 
percent.

5. The chance to gain and keep excess profits is essential as stimulus 
to entry.

6. If entry does not remain open after an initial entry, little change 
in industry behavior may result from the episode.

7. For entry to induce increased efficiency by prior firms, it must 
reduce profits so substantially that the prior firms must strain to 
reach “permitted” profit rates.

8. Imports are a major source of entry, but they are unavailable for 
most utility markets.

9. Where entrants are open to retaliatory entry by the utility, entry 
to the utility market will be inhibited and possibly excluded.

An entry fostering policy, therefore, will have small scope in 
vertically integrated, multimarket, and interrelated systems such as 
telephones. An effective policy to promote entry will aim at a series of 
entrants, whose eventual shares will constitute a large share of the 
market, gain large excess profits for a period, and continue to behave 
competitively in the long run. By contrast, a maximum of es-
suing in a tight, cooperative oligopoly still dominated by the original 
utility will yield little or no net benefits.

For the second type of entry — into the management function, by 
take-over or other means — these lessons hold:

1. The entrant must engage substantial support from outside finan-
cial sources.

2. The performance of the target management must have been 
clearly below a reasonable, attainable standard.

3. Regulatory and other authorities must avoid any devices by 
which the target management can prevent or delay the take-over.

Otherwise, the agencies will serve passively to exclude the take-
over. Such devices are already present in conventional regula-
tion.

Strategy to Optimize Entry

One theme of this essay has been that entry is likely to have little 
effect on utility behavior even if public policy toward entry is optimal; 
the other theme is that such a policy is likely to be a complex matter.
It is necessary to define the main elements of an “optimal” policy.
Although entry may at best be a minor factor, the changes required to 
optimize it may be small and inexpensive, so that the net social yield 
is high.

There should be no illusion that public policy can simply “permit” 
entry; instead, it will usually have to make entry. The very conditions 
of entry and survival, and the selection among potential entrants, 
will have to be set directly by regulators. They will continue to deal 
with specific firms, whose finances and technology will still have to be 
evaluated. There is no magic button marked “Entry” which, once 
pressed, will solve the big regulatory problems automatically. Only if 
regulation is radically revised will a free-entry strategy substitute 
for complex controls on profits and rate structure. Worse, a moderate
entry policy will make these controls more complex to apply, because more firms will be involved and the variety of behavioral outcomes—from shared monopoly to unstable cartel—will be richer. Even so, a realistic strategy to promote entry will include four main elements.

(1) **Open technology.** Access to technology is necessary if any entry is to occur. This requires not only open information, but also free interconnection of entrants with existing systems and ready availability of essential equipment. More subtly, it requires that the very design of utility systems maximize openness and separability. This is most obvious, for example, in telephones, where the integrated Bell system is under strong inducements to choose more exclusivity for its technology in all of these directions.82

Commissions seeking to foster entry must take positive action against these types of exclusivity. Rules requiring openness are necessary but not sufficient. In addition, the commission must understand the technology thoroughly enough to control against exclusivity in all major instances. It probably will also have to eliminate vertical integration on a monopoly basis, such as exists in the Bell and General Telephone systems. It will need to prohibit long-term contracting and noncompetitive arrangements which would make it difficult for entrants to obtain essential equipment. These steps require more resources and expertise in the commissions, not less. Presently, commissions virtually abstain from the subject. In the more complex cases (telecommunications, transport), a very large expansion would be needed in commission staff, powers of requiring information from carriers, and competence of commissioners.

(2) **Target market share.** An entry strategy will need to prepare for entrants' combined share of at least 25 percent and for preserving independence of behavior. Anything less will have insignificant effects, while adding to difficulty of regulation. The ceiling on market share will be the level at which the original firm is just barely unable to achieve "adequate" profits even while operating efficiently. Determining this level objectively usually will require sophisticated research.

(3) **Entrants' profits.** Excess profits will have to be permitted to entrants during a trial period. Later, an earnings constraint will

---


---

William Shepherd 59

usually be necessary, unless entry has proceeded so far — to essentially a moderate or loose oligopoly basis with four-firm concentration below 50 percent — that effective competition prevails. The proper length of the initial period, and the level of profits permitted later, will require, once more, sophisticated study.

(4) **Take-over rules.** To permit the take-over threat to act as a managerial stimulus, there will need to be drastic revisions in certain regulatory rules. The enforced separation of utilities from firms in other activities will need to be modified, both in state and federal regulation. As long as full and separate accounts are provided, the older abuses perhaps can be averted. This change is desirable also to increase the number and strength of potential entrants into the selling function. The 1910 Public Utilities Holding Company Act has turned out in practice to shield inefficient utility management.

There are many possible ways to facilitate take-overs, some of them fascinating and highly experimental. One is to hold periodic auctions at which outside groups could bid for managerial function.83 These could be extended to bidding for entire utility franchises (if partial entry is not feasible or forbidden). The process would favor the incumbents, who inevitably would have superior information and experience. Yet, some such bidding process would make it possible to replace seriously defective managers and to induce better performance. Alternatively, such bidding could be on an open-door basis. In any event, the proceeds could go to the public purse.

Another approach would be to permit any and all take-over attempts, provided only that the carrier's supply requirements would be contractually binding. All take-over groups would, of course, be portrayed by the target management as transient and unreliable, but this need not deter the commission from relying on normal contractual processes to function so as to ensure service.

There will need to be experimentation with these and other devices in order to learn which ones will work. Almost any marginal move away from the present total exclusion of take-over is desirable, and ultimately it may turn out that all restrictions against take-overs should go. At present, utility management are protected as no others in the economy, except banks, which may explain the love in which they are held among informed specialists. Commis-
sions cannot be expected to replace poor managers by direct action, but they could permit normal, informed commercial processes to do it.

If the argument here is at all correct, this sort of entry eventually may be more important than conventional entry (of new sellers) in most utility sectors. So far, the first type of entry has attracted much optimism, which may and should recede once the regulatory burdens imposed are realized. In fact, I find it difficult to specify any major utility in which the first type could be achieved to any substantial degree without a large expansion of regulatory effort. Entry will have to be created, with much artistry and care (more than is commonly shown by any regulatory commission now in being).

By contrast, the second type of entry could be fostered while reducing regulatory effort. It would improve only two dimensions of utility performance — static and dynamic efficiency — leaving other major problems for direct treatment, but at least it would not make unreal and unrealistic demands on a regulatory process which is already deficient in the quantity and quality of its resources.

IV. Entry into Power Markets

James R. Nelson

This chapter is concerned with the consequences for public policy when a hard shell of regulatory traditions is cracked by the pressure of growing technology.

The specific topic is "power markets," so evidence garnered elsewhere in the economy is useful only by analogy. Nonetheless, many of the issues which have surfaced only recently in the electric power industry have been recognized for a long time in other economic sectors. Since the economic effect of transmission of electric power is very similar to that of transportation of commodities, analogies from the transport area should be closer than others. Since transmission of electric power concerns a single product, the closest analogies of all should be found in the history of the development of pipeline transportation of petroleum (and its products) and of natural gas. But the transport sector also provides a number of general parallels to the problems now developing in the electric power industry: "Regulation" and "competition" have been expected to move in double harness for the railroads since 1897 and for other forms of internal transport since their federal regulation was first undertaken in 1935 or 1940. So the transport sector will be explored for further clues as to how competition and regulation may be made compatible.

The first two sections of this chapter will be concerned with the lessons of experience with regulating petroleum and natural gas...
pipelines. The remaining sections will broaden out to a number of the general questions which must be raised if the transmission of electricity is taken as equivalent to the transportation phase of what is now mainly a vertically integrated industry.

The latter part of the chapter will first examine the possibilities of enforcing a free entry policy of a type often approached by antitrust proceedings in the context of what are assumed properly to be competitive industries. Next, the problem of entry will be examined in its vertical as well as its horizontal aspects. Then will come the question of whether, and where, divestiture might be possible in the electric power industry in order to split off transmission from generation, on the one hand, and distribution, on the other. The argument proceeds to examine enforcement of common carrier status on transmission systems as an alternative to the possibilities previously examined. This possibility is narrowed, in turn, to the special kind of status determined for oil pipelines by the consent decree of 1941. Finally, these familiar questions of forcing hostility toward entry, or of regulating entry as free entry, or of special arrangements designed to make the industry act as if it were either competitive or regulated, are placed in the familiar economic framework of decreasing cost industries and in the special institutional framework which has been created in the electric power industry by the existence of publicly owned and cooperative suppliers who can obtain new capital more cheaply than investor-owned enterprises.

There are four underlying threads running through the entire discussion.

First, the electric power industry is often referred to as a "decreasing cost industry." The first question is: In a multistage industry, do "decreasing costs" exist at each stage; if they do, is the phrase "decreasing costs" given the same economic meaning at each stage? The next question is: Given the definitions of decreasing cost which must be employed at each level of the industry, what is the approximate economic importance of "decreasing costs" at each level of the industry relative to every other level? Second, the electric power industry is also often referred to as "a regulated monopoly." Yet, there are a large number of firms at all levels of the industry, and the total number of enterprises is by far the largest at retail, which is the level to which the "monopoly" label was first applied. Therefore, the word monopoly cannot be taken literally to refer to a single seller for the entire country. If it is restricted to local geographical areas, in an industry whose capacity for efficient transmission over long distance is steadily increasing, is there a possibility that a technologically obsolete definition of monopoly may be accepted as a basis for public policy?

Third, the electric power industry is, in the main, vertically integrated. This is especially true of the larger, investor-owned, segment of the industry, but it is by no means entirely true even of this segment. To the extent that vertical integration exists, does this in itself create a barrier to competition higher than those which would exist because of decreasing costs at specific levels of the industry? And are there any economies of vertical integration which may at least partially compensate for the economic effect of combining horizontal barriers in a vertical form?

Finally, given the extent and types of decreasing cost to be found at various levels of the electricity supply industry, and given the degree of horizontal and vertical integration now prevailing within the industry in the United States, can any realistic suggestions be made for changes in regulatory emphasis for introducing antitrust emphasis where none existed before? Can anything be done about the structure of the industry, and should anything be done?

The Petroleum Industry

Among all forms of transportation, the most relevant form for purposes of comparison with the electric power industry is the pipeline transportation of crude oil and its products.

Long distance transportation of crude oil, as does long distance transmission of electricity, involves very significant economies of scale as higher capacity facilities are installed. These economies are the source of the familiar tendency toward monopolization as the firms which expand first and fastest make the survival of their smaller and less efficient competitors more difficult and by their very expansion erect ever higher barriers to entry against all but the largest potential newcomers.1 At the same time, cheapening of inland transportation (or transmission) has been made easier as technology opens up new horizons of decreasing costs. This lowering of cost of transportation — or trans-

1 For a brief but authoritative description of the classic use of pipeline economics, by the elder John D. Rockefeller, to achieve and fortify a market position, see Arthur M. Johnson, Petroleum Pipelines and Public Policy (Cambridge, Mass.: Harvard University Press, 1967), chapter 1, pp. 5 – 19.
mission — broadens the geographical market accessible to any one producer at any one location and therefore improves the prospects for development of competitive markets.

The impact of this geographical flattening of the transportation cost structure tends to differ according to the conditions of production at point of origin. For crude oil, as for natural gas, these points are geographically concentrated relative to consumer markets, and they are subject to constant or even increasing costs once output expands beyond moderate levels relative to the size of the available market. For electricity, the attraction of cheap fuel has historically been offset by the negative effect of transmission costs. In the past, delivery of fuels close to consuming markets has generally been cheaper than on-site generation of electric power combined with long distance transmission. Moreover, the economics of scale which may be realized from the generation of electricity diminish relatively with increments of capacity beyond fairly low levels.

From the standpoint of economic policy, these differences create two areas of contrast. First, for crude oil and natural gas, the decreasing costs which accompany larger diameter pipelines tend to produce further geographical concentration of an industry which is already concentrated. This influence may prevail even if individual supply source is constant cost or increasing cost.

Second, for electricity, the effect of the even more sharply decreasing costs which accompany increase in transmission capacity is to increase possibilities for market overlap. This effect increases the potential geographical scope of competition. Reduction of transmission cost may shift generation of electricity toward fuel sources, but economies of scale in generation tend to pull generating plants toward the largest markets. Once nuclear power is thrown into the comparison, geographical localization is no longer determined by the location of the cheapest fossil fuels. Assuming that nuclear power practically equalizes generation costs from plants of equal size at different locations, then the entire effect of decreasing transmission costs is to enhance possibilities for potential competition. Whether these possibilities are ever realized, in fact, is a function both of the corporate structure of the electric power industry and of the regulatory framework in which the industry operates.

So much for preamble.

Crude oil pipelines are technologically and institutionally the nearest relatives of electrical transmission systems. Both provide the spinal columns for vertically integrated industries. Both provide significant economies of scale as capacity is increased per unit of distance. Both can have profound impacts on the geographical structure of competition, as scale economies reduce the geographical contours of energy costs to users.

But these similarities end with a contrast. The industry of producing, refining and transporting, and distributing petroleum products is presumed to be competitive; the industry of generating, transmitting, and distributing electricity is presumed to be monopolistic. Was the elder John D. Rockefeller correct in 1911, and the Supreme Court incorrect? Or is there something amiss with our current public policy toward the electric power industry? Or, for that matter, is the technology of the petroleum pipeline industry so strained, with respect to the electric power industry, that no benefit possibly can accrue from pursuing it?

As the preceding paragraph should indicate, any attempt to answer the question of how competitive a natural monopolist should be must become entwined with the answer to the opposite question: How monopolistic can one permit a supposedly competitive industry to become? And, as one struggles to disentangle both of these questions, the Laocoon group is completed by a final query: What does entry have to do with all of this, anyway?

It is not here that the analogy of the petroleum industry to the electricity supply industry may be most useful. Economists across the spectrum from Joe S. Bain to John Kenneth Galbraith have stressed the fundamental point that market organization is not merely a function of the interrelationships of firms already in an industry — with each other, with suppliers, and with consumers — but of the influence which may either be exerted, or not exerted, by the reserve army of potential competitors.

And it is exactly here that U.S. economic history has taken one of its more surprising detours. The history of public opposition to the old Standard Oil trust is a history of opposition not only to the process of buying out or driving out actual competitors, but also of driving out those who were driven out in such a way as to warn anyone who might otherwise choose to follow them. Conversely, in the history of the development of the electric power industry, the economies of scale present at all levels of the industry have been adduced, not as evidence of the need for extra vigilance in maintaining competition (à la the Standard Oil case of 1911), but as evidence of the need to restrain competition. We are led to the following paradox: If economies of scale are really that important, why try to retain competition? Conversely,
James Nelson

66

Entry into Power Markets

If economies of scale are even more important (as they certainly are, cumulatively, for all stages of the electric power industry), then why try to superimpose regulatory barriers to entry on the overwhelming technical barriers which would exist even without regulation?

The opposite policy conclusions derived from the similar technological facts of the "geographical" levels of the oil and electric power industry — involving long distance pipeline transportation, in one case, and long distance transmission of electric power, in the other — are probably themselves the products of argument by analogy. On the one hand, production of crude oil has always involved a substantial degree of competition. Refining and retail distribution have also been competitive. Conversely, the economies of scale relative to the size of the market served are greatest, for electricity, at the level of retail distribution. These economies also exist at the level of generation. So, if the same industry manages to pile three separate levels of "economies of scale" atop one another in one vertical structure, the idea of enforcing competition on the industry is chimerical. The problem is that an "industry" whose technique divides it into three, sharply distinct, vertical levels must be analyzed layer by layer. Decreasing costs at the stages of generation and transmission are not necessarily additive or multiplicative. The case for monopolization created by decreasing costs at the level of generation may be offset, not reinforced, by the case for monopolization created by decreasing costs at the level of transmission. Decreasing costs in transmission do not reinforce the monopolistic attributes of decreasing costs in generation. Instead, they make competition among generating plants possible despite possibilities for decreasing costs for generation at any one site.

The fact that the generation, transmission, and distribution of electric power under one ownership is more or less taken for granted is something more than historical curiosity. There is obviously nothing "natural" about it. The Central Electricity Board was established in Great Britain as long ago as 1926 to interpose a publicly owned national grid system between generating plants and distributors who had supply deficits. In the United States, municipal distribution without generation is practically as old as the industry, and the largest single producer of electric power in the United States, the Tennessee Valley Authority, operates primarily as a generation and transmission enterprise, not as a retailer.

So the vertical organization of the electric power industry which is often taken for granted can be explained neither by the logic of electrical technology nor by the uniform testimony of the present organization of electric supply undertakings. The explanation probably lies somewhere in the interplay of technology and history. Even after the triumph of alternating current, exogenous forces were reinforcing the original technical reasons for vertical integration. The public utility holding company movement was based on the electric power industry, and this movement had at least two attributes pointing toward vertical integration. The first was that holding companies sponsored by electrical equipment manufacturers were among the leaders in the field — and the largest U.S. manufacturers of electrical equipment were, and still are, producers of everything needed for the supply of electricity, from turbogenerators to light bulbs. Second, if anything could justify holding company pyramidling of leverage to investors, it had to be their possession of monopoly position in their markets. To be really impressive, this monopoly had to be vertical as well as horizontal. Insofar as this second reason, in turn, enabled the holding companies to achieve at least some economies of scale in generation and transmission which were not available to smaller enterprises, then a reciprocal relationship might be established: Horizontal expansion might aid in achieving economies at some other stage of the industry, and these might, in turn, provide a basis for outbidding other suitors for the hands of further distributing companies.

Finally, of course, geography can both alter the nature of the monopoly problem and play tricks with economic discussion of it.

Geography can alter the economic nature of the monopoly problem because the pre-1911 Standard Oil position of dominance in the U.S. market and heavy influence in the international market cast a much longer shadow, politically and socially as well as in terms of economic options foreclosed, than the "monopoly" possessed by the electric lighting department of some small municipality or even the largest of the pre-1895 holding companies. For residential consumers, a purely local monopoly of electricity distribution may be more costly or more vexatious than even the boldest attempted monopoly vendering some other product. But for large industrial buyers, there is every difference between a worldwide or national monopoly of the supply of an essential service, on the one hand, and keen competition among a number of different would-be sellers, on the other. A buyer who has potential geographical mobility, in the sense that he can choose the location of new plants without being confronted with marked comparative advantage for the seller at one location, can
practically avoid the effects of "monopoly" as long as the monopoly is local. He obviously cannot avoid a national or international monopoly. For obvious reasons of political backlash, the political rules may be just the opposite of the economic rules. The wrath of constituents is likely to have greater cumulative impact on congressmen if it is directed at only one "monopoly" enterprise, from all parts of the country, than if it is directed at a rather hypothetical monopoly category consisting, in fact, of a large number of firms with local monopoly positions but without national reputations.

As for the trick:42 that geography can play on the "monopoly" concept, the spatial elements in monopoly are familiar enough to economists, but often in a rather diluted and shifting way. The nature of the problem may be clarified by an appeal to economic history. The seemingly obvious tendency toward monopoly which was a major factor in the alarm which finally produced the Sherman Act may also be viewed as part of a defensive reaction to the increase in competition in major markets which was made possible by the rapid improvements in long distance transportation which took place after the Civil War. One effect of these improvements was to set the stage for differential advantages to large firms by way of differential railroad rates. Another effect was to place all kinds of firms economically neck-and-jowl when they had formerly enjoyed local preeminence, hence, there was a sudden increase in competition which might not have been as intense as the old level even after the elimination, by merger or bankruptcy, of some of the competitors.

We may conclude this section with the following summary:

1. Although the vertical integration of the electric power industry has a great deal in common with the vertical integration of the petroleum industry, and although long distance movement of energy (pipeline or by high-voltage transmission line) has been central to the development of competition in both cases, there is nothing quite so clear in the oil industry as this role of long distance transportation has been quite different.

2. In the case of the petroleum industry, the economies of scale in pipeline transportation created obvious advantages with respect to cost. The temptation to monopolize at the lowest levels of the industry. Recognition of these advantages provided the occasion, in turn, both for a range of regulatory powers to the Interstate Commerce Commission and for the 1941 consent decree entered into by the Department of Justice. The absence of any assent of competition at the retail level in electricity distribution, together with the fact that local generating plants had little opportunity to develop or demonstrate economies of scale in the early days of the industry, combined to create a policy vacuum until recently: Transmission lines were believed to be plant facilities purely and simply, without common carrier status and without an economic role in spreading the economies of scale in generation across ever wider geographical areas.

3. This acquisitiveness in the "vertical integration" approach to electricity supply cannot be sustained either by the technology of the industry or by the facts of its present industrial organization. In the United States, the answer must be sought not only in the relative world of downtown central stations and the trolley car, but also in the vertical bias introduced into the industry as a byproduct of the leapfrogging horizontalism of the public utility holding company movement. Enforcement of the Holding Company Act of 1935 has allowed geographically integrated, horizontal organizations to remain intact; the same has been true in effect for joint gas and electric companies; the same has been true, again — and without even a pretense of being raised except with respect to some local transit enterprises — for all forms of vertical integration. In this sense, the holding company of the 1920s is still very much with us.

4. This relative lack of interest in the vertical integration of the power industry, as compared with petroleum, may also be explained in part because monopolization in petroleum, if less intense even before 1911, has traditionally threatened to be more extensive. Petroleum pipelines stretch for very long distances, due to the geographical relationships of major producing (and refining) and consuming areas. Transmission of electricity for such distances has not been economical until very recently. Moreover, the effective market radius of an international oil company is far beyond anything found among distributors of electricity. The result has been a psychological and even legal preoccupation with the phrase "local monopoly" with respect to the supply of electricity, which tends to be at variance with the present-day technology of high-voltage transmission of electricity and hence tends to beg the most relevant questions of current policy both by using the word local and by using the word monopoly.

The Natural Gas Industry

The record of regulation in the natural gas pipeline industry should, in principle, be of more interest as a preliminary to an analysis of entry into the electric power industry than the record for
petroleum pipelines. At the producer level, much of the output of natural gas is controlled by the same firms that control the output of crude oil; indeed, a considerable fraction of total natural gas output is derived as a byproduct of the output of petroleum. Although the main production areas are not identical, the pattern is similar enough for purposes of this discussion. The one significant contrast arises from the fact that heavy tanker movements of petroleum and products from Gulf ports to northeastern cities have no counterpart in the natural gas industry. Where the industries differ, these differences would seem to favor placing most emphasis on a study of natural gas. Since 1945 the natural gas industry has grown much more rapidly than the petroleum industry, and practically all of this growth has been based on expansion of old pipelines and construction of new ones. As a corollary of this growth, the average distance between source of supply and consuming market has been greatly increased. In 1930, natural gas pipelines stretching out from producing areas in the midcontinent and southwest reached only to Denver, Omaha, St. Louis, and Atlanta; a dense system of lines in Ohio and western Pennsylvania was almost entirely confined to those two areas. By 1959, a continuous series of natural gas lines connected major southern and southwestern producing areas with points as remote as Miami, Concord, New Hampshire, Minneapolis-St. Paul, the Puget Sound area, and Los Angeles.

Thus, the stage would seem to have been set for a critical test, under comparable experimental conditions, of exactly the "entry" hypothesis being examined here. The possibilities for such a test were by no means neglected by the Federal Power Commission. Competitive applications were received, and at least quasi competitive decisions were handed down. In some areas, such as parts of New England, an acrimonious dispute in the best competitive tradition developed between rival claimants for the Federal Power Commission certificate of convenience and necessity. Any failures with respect to entry policy could scarcely be attributed to lack of interest on the part of the potential entrants themselves.

John R. Stockton, Richard C. Benshaw, Jr., and Richard W. Graves, Economics of Natural Gas in Texas (Austin: Bureau of Business Research, College of Business Administration, University of Texas, 1952), Figure 15, p. 156.

*Alfred M. Lasson, John A. Sibbett, and John C. Jaynes, "The Dynamic Natural Gas Industry" (Norman: University of Oklahoma Press, 1963); Figure 52, p. 275.

A great deal has been written about the Aggasiz-Northeastern imbroglio. See, for example, Leslie Cookerhbn, Jr., Crude Oil Pipe Lines and Competition in the Oil Industry (Cambridge, Mass.: Harvard University Press, 1950), pp. 147-49.

But for a number of reasons the natural gas case has never provided a clear-cut parallel to the situation being analyzed here. (1) One of the most important was precisely the dissatisfaction which prevailed during the late 1930s with respect to the success of efforts to control petroleum pipelines, either by encouraging competition or by Interstate Commerce Commission type regulation. As a present member of the Federal Power Commission describes the situation:

The major companies who developed the great new oil and gas fields of the midcontinent were the builders of the first long distance pipelines. The holding companies in the industry were the targets for the advancing pipelines, recognized what was happening, and jumped into the pipeline development activity. It was entirely predictable, therefore, that when governmental concern about concentration of economic power in too few hands, particularly in utility holding companies, reached fruition in the enactment of the Natural Gas Act of 1938, that the utility model or mold for such legislation was regulatory, and assigned to the FPC, already experienced in regulation under the Public Utility Holding Company Act of 1935.

Insofar as pipeline initiatives came from retail distributors of natural gas, a regulatory approach analogous to that employed to control final prices to consumers who purchased from these distributors naturally appeared to be in order. Insofar as pipeline initiatives came from large natural gas producers (who were also, in the main, large crude oil producers), disenchantment with the regulation of crude oil pipelines against a background of strong vertical integration would alone explain the interest in regulating prices and rates of return at the pipeline level itself.

(2) Effective federal control of pipelines had reached a critical interim stage during the period of greatest growth for the industry. The regulatory scope of the Federal Power Commission under the Natural Gas Act was not even spelled out, in general, until 1942 (by the Supreme Court decision in Natural Gas Pipeline Company of America v. Federal Power Commission, 315 U.S. 573); it was not until 1954 that the Court reversed the FPC and informed the commission that it was required to assume jurisdiction over prices at the producing and gathering level, for natural gas intended for interstate ship-

ment, and not merely sales prices at city gates (in Philips Petroleum Co. v. State of Wisconsin, 347 U.S. 672). Thus, during this critical, twelve-year period, the legal position appeared to be that fully integrated natural gas companies were subject to public utility type regulation based on their balance sheets and income statements for all phases of the industry, as opposed to vertically disintegrated natural gas companies which could apparently (until the Phillips decision) pay any prices they pleased for supplies of natural gas. Thus, enterprises which anticipated that they could obtain more profit from producing natural gas than from pipeline transportation of this gas, subject to standard public utility regulation, had what seemed at the same time to be a strong regulatory incentive to set up independent enterprises for production and pipeline transportation.

In 1945, gas utilities and pipelines produced 10 billion therms of gas and purchased 16.65 billion therms. By 1955, their own production had declined very slightly, to 9.95 billion therms, as contrasted with a fourfold increase in their purchases, to 65.4 billion therms. The pattern was set in this critical period. By 1969, integrated production had grown to 13.2 billion therms, but purchases had soared to 160.4 billion therms. The critical influence was apparently the FPC, before 1954, and not the Supreme Court in the Phillips case.

(3) The question of pricing at the producer level has come close to monopolizing the time the FPC devotes to natural gas. The Permian Basin has crowded out pipeline economics.

(4) The leading case which bears on the subject of this essay finally reached the Supreme Court as United States v. El Paso Natural Gas Co. (376 U.S. 651). The subsequent history of this case well illustrated the difficulty of generalizing from the specific attributes of an industry based on an exhaustible (and apparently rapidly depleting) natural resource. The increasing scarcity of natural gas has shifted the emphasis of many congressmen and state regulators from problems of competition to problems of availability. But, for all of its uniqueness, the natural gas record nevertheless has several contributions to make to the economics of the electric power industry.

First, for an industry involving significant decreasing costs, ver-


The Entry Options

The discussion so far has relied mainly on analogy. Henceforth, attention will be concentrated on the institutional characteristics of the supply of electricity. These characteristics can be summarized very briefly: The "representative firm" in the U.S. electric power industry is vertically integrated at least from generation through retail distribution, and in a few cases even farther back, to fuel supply.

As for public regulatory policy toward vertical integration, the FPC cannot disassemble or reassemble power systems. It cannot compel a vertically integrated enterprise to sell electricity, in the
industry is, de facto or de jure, a closed preserve, then at least common carrier status may be employed to require the kind of even-handed economic justice to users of the service of common carriers which facilitates entry at other levels.

(5) If common carrier status for the "transportation" level of the electric power industry appears to be inappropriate, the next possible remedy would seem to be of the "as of" variety embraced in the oil pipeline consent decree. This remedy involves an unusual amalgam of antitrust and regulation. A regulated industry which is at an intermediate stage in a chain of vertical integration is assumed, in standard regulatory fashion, to be capable of discriminating among different customers; it is furthermore assumed, in standard antitrust fashion, to be capable of independent operation if the regulatory approach to prices charged and profits earned appears not to produce beneficial social results. The industry is then treated as a possible orphan by antitrust standards. If the oil pipeline negotiations furnish a valid precedent, this potential orphan then buys his way back into the family by reaching an agreement with a nonregulatory body, which is similar to the kind of regulation which might have been applicable if, in fact, the appropriate regulatory body had regulated it.

(6) Whether or not approach (5) is pursued, there remains the central issue of how to facilitate access to decreasing cost facilities, which now stand in the middle of a vertically integrated corporate production line, on terms which are in some sense equivalent to those available to the upstream and downstream portions of the present integrated enterprise.

(7) Having used the word equivalent in place of the more obvious word equal in (6), one must then approach the bridge which separates electric utilities from petroleum. Farm cooperatives dabble in the refining of petroleum and are quite important in its resale. Otherwise, the oil industry is profit-minded from start to finish. But public bodies, and cooperatives, may borrow money for electrical enterprises at rates which are considerably below those available to investor-owned utilities. They need not pay income taxes on their own operations. And, if they are so minded, they may often fall back on a cushion of general tax revenues if their own projects go astray. In view of these various types of favoritism, why should an investor-owned utility ever give up its option to refuse any kind of participation with public bodies in order to prevent being overwhelmed, in the long run, by low risk and low cost competitors?
Finally, there remains the question of providing—or enforcing—incentives to deal with respect to sellers whose interests may be served as well or better by a boycott. For a seller whose overall rate of return is regulated, additional sales to any one type of consumer, and additional profit from these sales, may be alike of neutral value. These extra profits may be extra excess profits.

Entry at Face Value

The idea of taking entry at face value is the easiest of all the alternatives listed in the previous section to eliminate. As long as vertical integration remains the fashion in electric power supply, any hapless potential entrant would have to face a racecourse on which all hurdles have been stacked atop one another at the starting line. But if there are substantial entry barriers at each or several levels, vertical integration may create a difference which is not in degree but in kind. As indicated by the hurdle analogy, this difference may be described as due to the "stacking effect" and the "standing start effect."

Serial Entry

Once one moves away from the problem of simultaneous entry at all levels to entry, selection, at one level or another, the answer is not so easy. In an environment of decreasing costs at all levels, ability to proceed one level at a time should at least make the problem less staggeringly difficult, which leads to the next question: At which level is entry easiest? The quasi corollary of this question is: At which level would entry be likely to produce a gain (or the greatest gain) in economic efficiency? Unfortunately, there is no reason to believe that the real world answer to these two questions is the same.

This problem is best illustrated by starting the examination at the stage of retail distribution, where the analysis of "economies of scale" in the electric power industry originated. But it is here, also, that entry is economically least difficult. Mere villages have their own retail distribution systems for electricity. Given the character of regulation in many states, other villages may permit the current municipal franchise to expire and take over the supply of electricity at their pleasure.

The economics of village entry are reinforced by the public franchise of village entry. With municipal earnings from electricity supply exempt from federal income tax, a municipal body would indeed have to suffer from a poor credit rating to pay more for capital to be invested in its electricity undertaking than even the most highly regarded power company.

So the trouble is that the meaning of economies of scale, at the retail level of the industry, does not entirely dovetail with the meaning of what appears to be the same term at other levels. Electrical utilities started as local enterprises. They were given local monopolies, and they originally were regulated at the local level. All of these historical phenomena were the products of "neighborhood effects" in the most literal sense. Given a string of single family houses, each with a few light bulbs and without any other available way for consuming electricity, duplication of supply facilities would have been as undesirable as it was unlikely. Thus, at retail in this kind of environment, there could be no economic excuse for competition among supply enterprises in attempting to serve each residence. But, at the retail level, there would be no special premium for expanding sales beyond the volume that could be achieved in a fairly small community, which requires an investigation of what lies behind the retail level.

The immediate focus is on transmission, but this cannot be analyzed effectively without reference to generation. The first reason for this is that transmission, as a wholesale function, is not tied to neighborhoods or even communities as distribution has to be. The very existence of transmission networks is optional.

There are, in fact, at least three reasons for transmission systems as a matter purely of geography and not of the decreasing costs which result from increasing the capacity of transmission lines. The first can be illustrated by the example of hydroelectric systems. Remote generating sites may have comparative advantages even at low levels of capacity and output. Once it is economic to build transmission lines between such sites and consuming centers, then the decreasing costs which accompany higher transmission capacities may be sufficient to overcome the higher costs of generation at the favored sites. The second reason for transmission systems arises from an application of the principle of delivery factor at the wholesale level. Public utility supply of electricity could be economic even in the absence of all economies of scale in generation, transmission, and distribution, as a simple consequence of the application of the law of large numbers to the stochastic character of electricity demand. On a wholesale level, the mere difference between the coming of daylight and dusk at various locations, east to west, may provide enough diversity to warrant an expensive transmission line. The third reason
which accompanies larger generating units. This higher ratio may be reduced, in its turn, by access to a larger total generating system. But the resultant economy can be achieved only through much larger total capacity and represents an external economy whenever viewed from the standpoint of any one system. For each separate system to maintain adequate reserve capacity would be inefficient, for each separate system to maintain reserve capacity which would be adequate only with the existence of complete interconnection would be incompatible with supply responsibilities. Simultaneous achievement of both optimum efficiency and optimum reliability requires the application of the insurance principle which is a function of the size of any one generating pool, but of the total capacity of all generating facilities available to consumers at any one location via interconnection. The second difficulty with the power pool movement arises from the fact that it presents the public policy option of the frying pan and the fire. Firms which are independent at the transmission and distribution level are to combine to achieve economies of scale in generation, they must sacrifice some of their freedom of action if the joint investment is to be efficiently planned and operated. On the one hand, decision making is complicated by the need to rely on some kind of committee system which reflects the views of all participants in the pool. On the other, the fact of cooperation in generation is not obviously compatible with the competitive assumption of unfettered, individual pursuit of potential, large, new power consumers. Also, as with any other form of pooling, the need to make allowance for the kinds of divergence of interest which are inseparable from any degree of managerial independence is incompatible with the need for centralized and consistent long-range planning which is a necessary, if not sufficient, condition for optimum economy in providing generating facilities. If the optimum scale of electricity supply is larger than anything that can be achieved by power pools because the pool cost curve, as influenced by the number of participating enterprises, begins to rise, while the curve of generation-cost-transmission costs at a given level of reliability is still dropping, then the familiar principle of the lowest common denominator may come into play. If every extension of the largest retail supply area threatens some weakening of competition for supply to large potential consumers, and if power pools among independent retailers threaten diseconomies before they have rendered, or through separate control of prices charged at generating plants and at distribution step-down transformers. Given the nature of the decline of efficiency, on the other, may be simulated vertical disintegration. In U.S. practice, the traditional form into which this concept fits is that of common carrier status. Common Carrier Status: The Reality This section will be concerned with the reality of common carrier status, as distinct from the concept of the "as if" common carrier which will be analyzed in the next section. The reality of common carrier status need not be accompanied by the legal status of an intermediary (the "pure common carrier") as opposed to a principal (the "private transporter," who owns what he carries). The distinction can be clarified by reference to the petroleum industry, on the one hand, and the natural gas industry, on the other. These two products are often produced and shipped by the same companies, they sometimes come from the same wells, and they compete in a wide range of end uses. Yet, the regulation of pipelines which transport crude petroleum is under the jurisdiction of the ICC and is in principle identical with the kind of common carrier regulation applied to railroads, with the very important qualification that pipelines are not subject to the "Commerce Clause" of the Hepburn Act of 1906. Conversely, the regulation of interstate natural gas pipeline movements is entrusted to the FPC and operates on the basis of an assumption of purchase and resale of most of the natural gas transported. Legally, therefore, petroleum pipelines are public facilities in a sense that natural gas pipelines are not. As a matter of institutional organization and practical economics, the reverse is true: Petroleum pipelines are operated mainly as plant facilities, whereas natural gas pipelines are genuinely independent middlemen between field producers and distribution companies who buy the product at the city gate. Thus, regulation of the common carrier type may operate, de facto, either by the enforcement of charges for the service itself, or by regulation of payments to suppliers, and charges to customers, in such fashion as to produce what is deemed to be an appropriate value added by transportation. The applicability of this conclusion to the electric power industry is as follows: "Common carrier" status for transmission systems could be enforced, in practice, either through a "wheeling" requirement, combined with the approval of tariffs charged for the service, or through separate control of prices charged at generating plants and at distribution step-down transformers. Given the nature
of electricity, the familiar transport application of the term to refer to carriage of goods owned by others is impossible. No one is capable of tagging an individual kilowatt-ampere. But the idea of purchase and resale of equivalent quantities, with allowance for line losses, is perfectly feasible. With the advent of computerized control of load dispatching, even the instantaneous determination of marginal costs, at point of origin and point of delivery, presents no insuperable quantitative problem.

Thus, there is no fundamental institutional reason why common carrier status should not be required of major electric power transmission systems. The problem, rather, is that of a practical application of Occam’s razor. Granted that full common carrier status could be decreed and enforced, is this the simplest and most efficacious way to approach the problem?

The first step in answering this question is to point out that the origins of common carrier status, and its original application to railroads, were both largely related to problems which have no analogue in the electric power industry. The key idea of requiring a common carrier to transport everything, in all quantities is either irrelevant or misleading when applied to the electric power industry. The further idea that common carriers should serve all feasible destinations cannot be thrown out as easily in the case of electric power industry, but it still runs into formidable practical problems with respect to the issue that has plagued the attempt to regulate petroleum pipelines as common carriers: How does one lend installations which were originally intended as plant facilities in the direction of general public service? The fact that origins and especially designated services by electric power transmission lines may be geographically much more diverse and scattered does not ease the difficulty. In some ways, it makes it worse.

Let us say that a distributor of electricity desires to receive bulk supplies from a lower cost source in place of its own inefficient generation. Let us suppose, further, that potential competition to render such supplies, in the form of major transmission lines passing near this isolated distributor, is somewhat on the scale between true monopoly and anything that could be trusted, market-wise, as acceptable oligopoly. Then an attempt to impose common carrier status on one possible supplier would be clearly discriminatory, whereas an attempt to impose such status on all possible suppliers would immediately raise further questions: Who is to pay for the construction of extra transmission lines to make the potential common carrier competition effective at destination? Who is to pay for shifting facilities from one site to another, or abandoning some facilities and duplicating installations at other locations, as the buyer exercises his option under the common carrier concept and shifts his source of supply? What is to be done about the uncertainty involved, for buyer or seller or both, in a situation in which common carriers may be required to plan ahead for markets which they may or may not be able to hold?

These questions are merely indicative of the thorny issues which would necessarily attend the application of traditional “common carrier” thinking to the electric power transmission industry. They should be sufficiently to indicate that the trouble with common carrier regulation of pipelines was not simply nonfeasance on the part of the ICC. As long as an industry remains vertically integrated, major problems must arise from any attempt to open the bloodstream and make it available to the public “without preference or prejudice.” They are great enough to repel the traditional common carrier approach to second-best status, or lower.

Common Carrier Status: The Concept

This brings us to the als ob equivalent to common carrier status. Can electric transmission lines be treated like oil pipelines, in the form of a limitation on earnings at this level of the industry in order to take the excess profit out of discrimination against outsiders?

It must be admitted, at the outset, that the legal backgrounds are very different. The Department of Justice was able to move against the oil companies under a rare statute, the Hepburn Act, which gave it jurisdiction over some transportation issues coordinately with the ICC. It was able to move, specifically, because the jurisdiction permitted it to challenge excessive pipeline profits as in effect constituting deferred rebates. Even with these special advantages, or perhaps because of them, the intervention of the Department of Justice produced no definitive judicial decision which could be used with respect to electric transmission systems or other such geographical interdependencies; the case was settled by consent decree.

But the practical case for the pipeline type of consent decree is probably a good deal weaker in the pipeline industry than in the electric power industry, where it has not been tried, and where, indeed, no present legal basis for an attempt would seem to exist. All earnings of electric power suppliers (or at least of investor-owned
suppliers, who control the major transmission systems except in the Tennessee Valley and parts of the West, are subject to effective control. Moreover, it can scarcely be argued that sources of electricity supply are temporary, shifting, and uncertain in nature. Vertical integration has apparently created a seamless web from generating plant to customer's light switch. There are parts of the country where transmission systems have also created a seamless geographical web at this level of the industry. The only difference is that the vertical relationship is usually controlled by one owner, whereas the horizontal relationship is not.

A further consideration is rapidly growing in importance. The siting of both transmission systems and the generating plants which they connect with markets for electricity has begun to come under critical review for environmental reasons. Until recently, no one seemed to care where a generating plant was located, what kind of fuel it used, or how it was connected to its markets. Indeed, both generating plants and transmission systems might be sought after as sources of growth in the tax base unmatched by comparable growth in welfare rolls or school populations. Economists and commissioners were wont to applaud "promotional rates," which cut prices so as to maximize growth of consumption.

The abrupt shift in opinion about the relative desirability of cheap electricity versus high-visibility transmission lines (or high-sulphur fuel) has been combined with a whole series of events which seem to foreshadow the end of cheap energy in the United States. Therefore, we can no longer blithely entrust most regulation of electric power suppliers to state bodies and muddle through at the federal level with overlapping jurisdictions, and conflicting objectives, and without overriding direction of policy.

Granted the emergence of a new framework within which all energy issues will in future have to be examined and settled, how does this affect the immediate question?

Possibly in this way: All regulation of transmission of electricity should be centralized at the federal level. This regulation should include both affirmative and negative powers over location of generating plants, as well as of the fuel they consume; it should include powers to determine the routes of transmission lines and to specify whether transmission is to be overhead or underground; and it would probably have to include affirmative powers, to compel required expansion, as well as traditional negative powers to prevent it. Regulation, whether federal or state, may simply have to deny the right to generate to certain electric power suppliers, including suppliers who now generate for themselves. This denial would have to be accompanied by provision of alternative sources of supply, which would, in turn, require regulatory control over what these sources charge as well as regulatory control of where they are located and how they transmit the product.

Thus, we reach the fundamental point and the fundamental paradox: Developments in the energy sector are bound to produce a rapid growth of controls over a number of the usual accompaniments of freedom of entry, such as free choice of location, free choice of raw materials, and free choice of distribution systems. Since these de jure controls will descend on a sector which is already characterized, in the electric power subsector, by de facto vertical monopoly, "as if" standards must be applied to make the transmission stage of the electric power industry behave as if it were part of the energy sector as a whole and not just an intermediate stage in the distribution of electricity. These "as if" standards will be greatly facilitated because of the long history of control of rate structures and rates of return in the electric power industry.

It must be admitted, as a bridge between this section and the next, that "as if" seems to be the little man who wasn't there. The next section will confront the more detailed issues presented by the suppositional character of this concept. But, in closing, it should be emphasized that the economic attributes and performance of what is deemed to be beneficial in the concept "as if" may be approached with common carrier status in its pure, old-fashioned form. The general problems which these policies must confront are introduced in the next section.

The Problems

A policy of "as if" is a policy hanging in the air. It fails to go all the way in answering the critical question. "As if what?" The obvious answer, "as if the electric power industry, at all its stages, were a competitive industry," or "a workably competitive industry," or "a purely competitive industry," only opens the door to a whole new series of questions.

(1) From the standpoint of rates of return, the "as if" question assumes the guise of "comparable earnings." To the extent that an
industry with sharply decreasing costs at all levels has built-in protection against entry, the existence of decreasing costs for technical reasons should be reinforced by decreased unit financial costs (that is, a lower "normal" rate of return). If this built-in protection is further buttressed by regulatory protection, the cumulative effect of the two influences should be to lower comparable earnings. If the two forms of protection are combined to produce market position which warrants raising an abnormally high proportion of required capital by the sale of debt securities, then the legal status of interest payments as a cost which may be subtracted before calculation of "income" for tax purposes adds a further institutionalized reason for differentiating between electric power earnings before taxes and other earnings before taxes, as a basis for determining comparable returns. Thus, electric power companies have been able to enjoy higher credit ratings than industrial trials and (until recently) comparable or higher price-earnings ratios on common stocks, despite long-term debt loads equal to 50 percent or more of total capitalization.

(2) Assume that comparable earnings come spilling out of some magic computer. Assume that this sacred earnings percentage need not be corrected, absolutely or relative to returns in other industries, from now on, forever more. There would still remain the separate problem of rate structure: What configuration of rates, amid the theoretically eligible candidates, should be chosen over all others?

For a purely competitive firm, the idea of "configuration of rates" is at least as pointless as the idea of "comparable earnings." The firm earns what it can get by charging the only price it can charge. But the same decreasing cost ingredients which operate to reduce the level of required utility earnings, relative to other industries, also operate to create a spectrum of average prices and another spectrum of marginal prices by way of which these earnings may be achieved. The usual way to approach such a spectrum problem is to go beyond pure cost of service considerations of value of service. Yet, "what the traffic will bear" does not always provoke a single number. A source of variability in what the electrical traffic will bear is the fact that opportunity costs associated with service employment are a prime determinant of service. Opportunity costs are themselves likely to be influenced by decreasing cost elements in the means that would have to be employed to provide service by alternative methods. When both options in a trade off are influenced by decreasing costs, then relative marginal cost coefficients may depend in part upon relative total quantities.

When these two points are combined, the conclusion must be that the purely competitive model fails short of providing a recipe for pricing on two counts: It produces the wrong target rate of return, and it has nothing to offer with respect to price structure.

A Possible Solution

It is precisely here that market defects may be introduced as a partial remedy for the others.

The first hypothetical defect has already been described. A large electric utility may be assumed to possess generating plants which already achieve scale economies not yet available to next-door neighbors; moreover, this large utility should also have realized new economies of scale in transmission. If the rate of return of this utility is rigidly controlled, the whole problem of the prices it should be allowed to charge for wholesale power, or for wheeling, may seem to be past all solution. Average cost pricing will carry rates of return above the permitted level; but if returns are regulated very severely, the attitude of the large power company may be capricious with respect to its own customers as well as outsiders. (Why be ingenious, for the sake of a zero net return?) A high price which fails to obtain outside custom may be more desirable than a low price which does obtain it, but at the risk of loss of control over the response of profitability to load growth. However, if returns are regulated slackly, there is no guarantee that the firm's pricing policy will either attract wholesale business even when it would be economic to do so, or pass on the unit savings derived from such wholesale business to retail customers of the integrated supplier.

The second defect is familiar and much publicized. It arises from the fact that the fiscal and financial position of investor-owned suppliers of electricity is inferior to that of governmentally owned suppliers or cooperatives. Investors have retained a dominant position in the industry despite their disadvantages with respect to raising capital in a capital-intensive industry due to two circumstances. First, hydroelectric power sites, which provided the original spawning grounds for large-scale utility, have now entered into a stage of sharply increasing costs. Second, investor-owned utilities have had the advantages of scale, at both the generation and transmission levels of the industry, outside of areas originally preempted by large public hydro projects.

Once these two wrongs are in confrontation, there emerges a not-wrong which may be close to right. If as is indeed the case, most
James Nelson 89

tegrated. If the institutional structure of the industry were frozen forever, these charges would deserve the utmost consideration. In a private enterprise economy, why should investor-owned utilities be further discriminated against when they attempt to escape from the first discrimination by purchasing power instead of generating it? Moreover, if these investor-owned distribution utilities were an important feature of the industry, then the paradoxical answer to this problem might have to be public ownership of the transmission systems to which they are connected. But the nonintegrated investor-owned utility is in fact a vestigial remnant. The present organization of the investor-owned segment of the electric power industry leaves little room for isolated distribution companies and not much room for really small integrated enterprises. Their possible complaints of discrimination could best be handled by encouragement of merger with larger companies, or by municipalization.

It turns out, then, that the "as if" approach to the problem of entry into the electric power industry resolves itself into regulating wholesale power rates and charges for wheeling power, so as to provide vertically integrated utilities and municipal distribution with lower cost power whenever this is possible, that is, by using the standard of marginally better off, for each, as the upper and lower boundaries of the zone of reasonableness.

Conclusion

There remains the key question with respect to anything resembling common carrier status: How can owners of transmission systems be forced to make their facilities available to others, if they choose not to make an outright sale and have no interest in providing wheeling services voluntarily?

From the standpoint of economics, this problem should be less difficult with respect to the transmission of electricity than with respect to most common carrier services. Railroadsundeniably carry large volumes of freight at a loss, by any definition of loss. The principle of averaging has deeply penetrated general utility rate structures. Against this background of de facto cross-subsidization, the sharply decreasing costs which attend growth in capacity of transmission systems should provide considerable assurances that the proceeds of marginal sales will not unduly depress average receipts. This conclusion is too straightforward in one way. A transmission system does not exist simply for the one-way movement of electricity,
as a pipeline exists for the one-way movement of petroleum. Therefore, transmission services cannot be accurately measured by expressing them in terms of kilowatt-hour or kilowatt-ampere miles from A to B. Insofar as capacity is an independent parameter of performance, the extra dimension can be managed as it always has been in the electricity supply industry: by use of two-part tariffs containing demand or capacity charges as well as energy charges. But the general stand-by services of a transmission line, and more particularly the general improvement in reliability provided by a transmission system, cannot be measured by any unit of volume. However, even this insurance role can be expressed, for ratemaking purposes, as some function of the wheeling charge for firm power. Indeed, to the extent that a transmission system contains elements which would be redundant except for insurance purposes, a cost basis for ratemaking which picked up some of these elements would automatically be reflecting insurance functions as well as direct service functions.

From the standpoint of the history of public utility regulation, the issue is not so simple. "Transmission" is simply one type of transportation, and the economic role of transmission of electrical energy is similar to the economic role of the transportation of tangible forms of energy such as coal and petroleum. But the history of regulation of transmission has been entirely different. The disinterest of the ICC in reinforcing common carrier responsibilities on petroleum pipelines for a generation after 1906 does not provide an entirely valid analogy, because the ICC was clearly interested in protecting and preserving rail rates and volume in the movement of petroleum. But the ICC at least had developed a clear concept of the meaning of "common carrier" and a great deal of expertise in applying this meaning to particular cases. No regulatory body concerned with public utilities has had comparable experience. It is entirely possible, then, that vertical disintegration of the electric power industry may have to be faced as an alternative to regulation, just as "yardstick" competition was once introduced as an alternative to regulation. Meanwhile, the situation requires much more than a mere regulatory change of heart, or a few Supreme Court decisions. For the moment, the groundswell of public opinion is to be found along the coastlines of environment and ecology. The emphasis on cheapness of electricity which was so prominent until recently now seems to be submerged beneath the ecological waves. But it will return; and it is likely to return with redoubled emphasis precisely because the high tides of the environmental movement are likely to push electricity prices far up on the beach.

Thus, the important task for the present is to begin to rethink our entire approach to traditional regulatory principles in the assumed context of the next round of regulatory dilemmas: when it turns out that expensive energy is an element in our balance-of-payments problem, or is a key factor in the consumer price index, or is interfering with the growth of productivity per man-hour. The response to these problems is likely to be not only a renewed interest in cheapness, as such, but also a brand-new interest in spreading this cheapness as far as possible geographically, in view of the desirability of strict control of the location of generating plants and the further desirability of trying to achieve a more rational trade-off between the security and capacity provided by transmission lines, on the one hand, and the environmental damage such lines may create, on the other. A return to the ideological battling of the 1930s would be a retrograde step. What is needed is adequate and nondiscriminatory access of distributors — companies, municipalities, and cooperatives — to major transmission systems and to the economies of scale which they possess.
V. Entry Conditions in Telecommunications

Harry M. Trebing
and
William H. Melody

The removal of the legal barriers to entry in the field of public utilities has attracted considerable attention in the past decade as a workable alternative to the seemingly insurmountable problems associated with commission regulation. An examination of the common carrier communications industry provides an excellent illustration of the effects of partial removal of entry restrictions and the consequent range of issues that are associated with the entry problem. In 1968 – 1969, in two landmark decisions, the Federal Communications Commission struck down tariff provisions which banned the use of foreign attachments on the telephone network¹ and permitted the entry of Microwave Communications, Inc., into private line communications. The first decision (Carterfone) promoted the entry of

new equipment suppliers and the so-called interconnect companies; the second (MCI) promoted the entry of specialized carriers into the private line market.

People may argue that these cases are not relevant, since legal barriers to entry remain solidly in place in message toll telephone and exchange telephone services, and these services account for the overwhelming share of the total revenues of common carriers. Still others may argue that the competitive entrants are not subject to the traditional obligations of public utility status, as are the common carriers, and therefore meaningful comparisons are not possible.

Nevertheless, the move toward liberalized entry has done more to create controversy in the communications field than any other single factor at the public policy level. The FCC has been placed in direct confrontation with state commissions over the issue. Pressure groups and lobbyists have emerged to represent equipment suppliers, specialized carriers, and interconnect companies. Established telephone carriers have mobilized substantial resources to resist entrants and mount counterattacks on a variety of fronts, and AT&T, the dominant firm in the industry, has called for a moratorium on all competition in communications. 8

This chapter will seek to evaluate the liberalized entry experiment. It will focus on the interrelationship between entry policies, industry structure, and corporate strategies. It will also delineate the barriers that impinge upon the entry process and examine the unique role of regulatory constraints. Finally, it will explore several options for public policy and present a summary assessment of the role of liberalized entry in communications.

Overview of Industry Structure

It must be emphasized at the outset that the changes in entry policy promulgated by the FCC involved the relaxation of selected legal barriers to entry. There were no concurrent changes in the general structure of the industry, and no changes in regulatory format or basic objectives. It is important, therefore, to appreciate the organization of the industry and the regulatory setting, for these factors exercise a pervasive influence on the entry process.

The Bell System, the independent telephone companies, and the so-called interconnect companies; the second (MCI) promoted the entry of specialized carriers into the private line market.

People may argue that these cases are not relevant, since legal barriers to entry remain solidly in place in message toll telephone and exchange telephone services, and these services account for the overwhelming share of the total revenues of common carriers. Still others may argue that the competitive entrants are not subject to the traditional obligations of public utility status, as are the common carriers, and therefore meaningful comparisons are not possible.

Nevertheless, the move toward liberalized entry has done more to create controversy in the communications field than any other single factor at the public policy level. The FCC has been placed in direct confrontation with state commissions over the issue. Pressure groups and lobbyists have emerged to represent equipment suppliers, specialized carriers, and interconnect companies. Established telephone carriers have mobilized substantial resources to resist entrants and mount counterattacks on a variety of fronts, and AT&T, the dominant firm in the industry, has called for a moratorium on all competition in communications. 8

This chapter will seek to evaluate the liberalized entry experiment. It will focus on the interrelationship between entry policies, industry structure, and corporate strategies. It will also delineate the barriers that impinge upon the entry process and examine the unique role of regulatory constraints. Finally, it will explore several options for public policy and present a summary assessment of the role of liberalized entry in communications.

Overview of Industry Structure

It must be emphasized at the outset that the changes in entry policy promulgated by the FCC involved the relaxation of selected legal barriers to entry. There were no concurrent changes in the general structure of the industry, and no changes in regulatory format or basic objectives. It is important, therefore, to appreciate the organization of the industry and the regulatory setting, for these factors exercise a pervasive influence on the entry process.

The Bell System, the independent telephone companies, and

Harry Trebing and William Melody 95

Western Union provide virtually all of the nation's common carrier telephone and telegraph service. Unlike the electric and gas utilities, the domestic communications industry is dominated by the holding company form of organization. The parent holding companies provide a centralized control function for a comprehensive network of operating companies as well as for vertically integrated manufacturing and research affiliates. As a result, the holding company typically embraces both communications services and communications equipment.

Within this setting, the dominance of AT&T is overwhelmingly apparent. Bell provides 83 percent of the exchange telephone service (measured in terms of number of telephones in use) and almost all of the plant and facilities dedicated to long distance or message toll telephone service. Its manufacturing subsidiary, Western Electric, accounts for approximately 85 percent of the sales in the domestic equipment market. The remaining 17 percent of the exchange telephone market is supplied by independent companies, but again holding company organization, concentration, and vertical integration are common characteristics. As evidence of growing concentration, the total number of independent companies has declined dramatically over time (from 6,400 in 1940 to 1,843 in 1970). At the present time, 25 holding companies account for 82 percent of the stations in the non-Bell exchange market. As in the case of AT&T, vertical integration backward into equipment supply is a common feature among the largest independents.

It should be noted that new entrants in communications must deal with these holding company systems. The specialized carrier (as a new entrant in the private line market) must secure local loops in order to reach the customers' premises and offer complete point-to-point service. The interconnect company (as a new entrant in the terminal and station equipment market) must be able to supply a product or service that is readily attached to the common carrier network. Clearly, communications remains a highly interdependent and integrated industry in terms of both the physical provision of service and the development and marketing of new offerings.

The market structure of communications can also be viewed in terms of the overall distribution of AT&T revenues. At present, the firm's interstate revenue constitutes 30 percent of its total revenue receipts. Approximately 87 percent of this interstate revenue comes from message toll telephone and wide area telephone service (WATS), and these amounts are growing at an average rate of ap-
proximately 12 percent per year. AT&T's private line, program transmission, and related specialized services (which will be grouped for purposes of discussion in the private line category) account for the balance of interstate revenues, or approximately 13 percent. However, private line revenues account for only 2–4 percent of total Bell System revenues. Thus, the so-called competitive private line markets constitute only a minute share of the present total. Interconnect markets are also relatively small, but accurate revenue estimates that would permit a better evaluation of size are not readily available.

The composition of future markets is, of course, a crucial factor. Forecasts require estimates of both the voice market and the new markets associated with data, record, and video communications. Although the growth rate in the voice market has been remarkably stable, the new markets involve a host of imponderables that make individual projections vary widely. This indeterminacy undoubtedly makes these markets even more attractive to both new entrants and common carriers, particularly if the established carriers view the traditional telephone market as approaching saturation in terms of the percentage of households with basic service. AT&T has moved to supply hardware and terminal facilities for portions of these new markets, and it has also designed rate structures which will attract them to the Bell network. But Bell has not moved to provide data processing services directly to the public. The independents (primarily General Telephone and Elecronics and United Telecommunications) and Western Union have promoted data processing services through their affiliates.

**Regulatory Setting**

The formal structure of social control in communications is composed of an admixture of state and federal agencies. State commission regulation began in 1967 in Wisconsin and New York; federal regulation began with passage of the Mann-Elkins Act in 1910, bringing interstate telephone and telegraph communications under the jurisdiction of the Interstate Commerce Commission. Since that date,

2. For example, the earlier forecasts of the growth of data transmission, see J. R. Pierce, "The Transmission of Computer Data," *Scientific American* 215 (September 1966): 148. Also, *Communications News* 7 (December 1970): 19.

Harry Trebing and William Melody

most state commissions have been given authority over intrastate exchange and intrastate toll telephone service, while in 1934 the FCC assumed and broadened the interstate responsibilities given earlier to the ICC.

Over time, a complex set of regulatory guidelines and ratemaking practices has emerged. At the state level, attention has focused on the control of the level of earnings for exchange and intrastate toll service within the context of the principle of statewide ratemaking. Under this approach, total revenue requirements are developed for the entire state, and revenue assessments against particular exchanges and services are then determined on a value-of-service basis. As a result, the largest exchanges are expected to charge a higher rate per station than the smaller exchanges, and vertical services (PBXs, key telephone sets, color phones, extensions, and so forth) are expected to yield a higher return than basic telephone services.

At the federal level, control of earnings, value-of-service pricing, and system-wide averaging became cardinal tenets of ratemaking. Most regulatory attention focused on control of the level of earnings, while MTT and private line pricing were premised on system-wide averaging, uniformity of mileage blocks with distance, and ample recognition of value-of-service considerations. It was in the switched message telegraph and private line telegraph markets that attention first focused on the interrelationship between pricing and industry structure. An uneasy duopoly existed in these markets, and the FCC apparently recognized the asymmetrical distribution of corporate power between AT&T and Western Union. The commission required fully distributed cost studies for TWX in 1934 and 1952 and for private line telegraph service in 1955. The results indicated relatively low levels of earnings on AT&T's competitive telegraph services.

Beginning in the late 1950s, Bell's pricing practices took cognizance of the impact of new technology on market structure. Providing a possible loss of business to new entrants or to customer-provided facilities, Bell filed a series of new tariffs in rapid succession. In 1960–1961, the FCC found itself confronted by WATS, WADS, and Telpak

In 1919 the Wisconsin Commission did attempt to pursue a cost-of-service approach to exchange ratemaking. But this effort collapsed in the face of carrier opposition. See *In Re Cross Telephone Co.*, PUR 1916 A, 522, and *Boyert v. Wisconsin Telephone Co.*, PUR 1916 C, 1029.

On occasion the control of earnings has been the causal factor influencing a change in rate structure. This was particularly evident when the need to reduce earnings realized in the introduction of "after-9" off-peak pricing.

In 1919 the Wisconsin Commission did attempt to pursue a cost-of-service approach to exchange ratemaking. But this effort collapsed in the face of carrier opposition. See *In Re Cross Telephone Co.*, PUR 1916 A, 552, and *Boyert v. Wisconsin Telephone Co.*, PUR 1916 C, 1029.

On occasion the control of earnings has been the causal factor influencing a change in rate structure. This was particularly evident when the need to reduce earnings realized in the introduction of "after-9" off-peak pricing.
Entry in Telecommunications

A, B, C, and D. The commission proceeded on a case-by-case basis, holding WATS to be lawful but eliminating WADES and Telpak A and B. The FCC then found that eliminating a portion of the Telpak tariff posed a further problem in the form of Telpak sharing, with clear implications for arbitrage and entry.

The growing problem of cross-subsidization and revenue accountability by class of service was dramatized by the 7-Way Cost Study in 1985. There was a clear need for new pricing guidelines; however, the FCC’s continuing search for such guidelines in the intervening years has not produced tangible results. In the meantime, there has been a new proliferation of pricing proposals and tariffs from AT&T. These pricing strategies will be discussed later, but it is sufficient to note at this point that pricing and entry are inexorably tied together. Regrettably, the FCC has been unable to delineate this relationship in its decisions, and the development of pricing guidelines continues to elude regulatory grasp.

Entry Policies

In contrast to the history of regulation which has prevailed in the area of pricing guidelines, federal regulatory policy toward entry has moved consistently toward greater liberalization.

Until 1959, both state commissions and the FCC followed the same general pattern. A relaxation of entry barriers required a showing by the potential entrant that existing service was inadequate or inefficient and that new entry was the appropriate procedure to improve service. When this policy was augmented by requiring the prospective entrant to demonstrate adequate financial resources as well as an assured customer base, it became a barrier to entry every bit as powerful as the neoclassical natural monopoly model.

The FCC began to depart from this traditional policy in the “Above 800 Megacycles” case in 1959. Although the new posture toward entry was formulated on a case-by-case basis, a common

theme recurs: The relaxation of legal entry barriers was deemed appropriate to promote the public interest by giving the consumer greater freedom of choice.

As noted earlier, the Carterfone decision (1968) marked a major step forward in relaxing entry barriers. The FCC held that the argument of systemic integrity was not controlling, and that Bell’s foreign attachment tariff had never been sanctioned as a lawful practice. The commission ordered the Bell affiliate to connect with a private mobile radio system, and thereby opened the door to potential entrants seeking to supply station and terminal equipment as well as other interconnect services.

The MCI decision (1969) authorized point-to-point private line service by a specialized carrier and brushed aside arguments which relied on cream skimming, lack of need, and the applicant’s financial resources and technical competence. In May 1971 the FCC further strengthened its position on liberalized entry by a general policy declaration in favor of open entry by specialized carriers.23 In domestic satellites, a similar “open skies” policy was adopted.24

The FCC has subsequently moved to restrict the ability of common carriers to circumvent or negate the new entrant’s ability to compete. In the specialized carrier field, the FCC ordered AT&T, on 23 April 1974, to provide the interconnection necessary for MCI and other specialized carriers to offer foreign exchange service and common control switching arrangements as a part of private line service.25 The order further directed AT&T to cease and desist from practices which delay interconnection for the specialized carriers and which deny services to these carriers on a par with AT&T’s own private line supplier, the Long Lines Department. This order was sustained by the U.S. Court of Appeals on 11 September 1974.26

Finally, the FCC has sought to delineate a “boundary” between the regulated common carrier services and nonregulated data pro-

---


Harry Trebing and William Melody

99
subsidiaries. As part of this estimate a forecast of potential market penetration was made for interconnect and private line services. The interconnect study focused primarily on PBX and key telephone service systems. The total US X service was assumed to grow at an average annual rate of 10 percent for the period 1974 through 1984. Key telephone systems were assumed to grow at 10.6 percent for the same period. Interconnect company supplied PBX services were also assumed to grow at 20 percent minimum average annual rate, while interconnect services were assumed to grow at 35 percent. Given these assumptions, Bell's share of the PBX market would decline from 84 percent in 1974 to 61 percent in 1984, and its share of the KTS market would decline from 99 percent to 85 percent in the same period. The interconnect companies' share would increase by 1984 to 35 percent in PBX and 15 percent in KTS. Specialized carrier diversification, and therefore market penetration, was not estimated in a fashion that is readily susceptible to the market share growth attributable to the specialized carriers, but it is apparently assumed to be significant.

However, there are at least three major reasons for doubting that entry in communications will have the serious impact on market shares of the telephone carriers that the NARUC committee anticipates.

First, the committee's estimates are based on limited information, supplied primarily by AT&T, and in order to forecast it was necessary to make a number of heroic assumptions and accept a great deal of averaging.

Second, the actual growth patterns of AT&T interconnect services indicate no discernible shift in intercarrier preferences that would sustain a belief that large shifts will take place in the future. The remaining constant. If one compares the five-year period prior to Carterfone (1963 - 1967) with the five-year period after Carterfone (1969 - 1973), the average annual rate of growth in revenues for all AT&T vertical services has remained a constant 10.1 percent.


"Before the Federal Communications Commission In the Matter of Applications of Telephone Companies for Section 214 Certificates for Channel Facilities Furnished to Affiliated Community Antenna Television Systems, Docket No. 16,909, Final Report and Order, 28 January 1979."
most aggressive rivalry between Bell and the interconnect companies has been in PBX equipment, and in this area the average annual rate of growth of Bell revenues has increased from 8.8 percent in the pre-Carterfone period to 11.1 percent in the post-Carterfone period.

Table 1 sets forth this information in greater detail.

Third, the experience of firms competing in the specialized carrier and interconnect markets has not been impressive. A number of the original entrants have seen absorbed or have disappeared through merger. For example, Southern Pacific Communications absorbed Video Microwave and United Video, Inc., in 1972. The latter had nearly completed construction of a specialized voice and telecommunications network serving the Southwest.14 General Dynamics acquired the well-known Aratoa Interconnect Companies in 1973,15 while MCI and Nebraska Consolidated Communications Company merged in 1974. Plans were also in progress in 1974 for MCI and Western Telecommunications, Inc., to integrate their facilities.16

Other firms have left the market or have experienced financial difficulties. General Electric withdrew from the PBX business in 1973,17 and Litton Industries closed its direct sales office and terminated employment of its sales personnel handling its business telephone systems equipment in 1974.18 Litton turned instead to independent distributors to market its products. Others, such as Phone-Mate (which manufactures answering devices and dialers) reported losses in 1974.19 Datran reported that it was compelled to seek capital from a Swiss financier because this was the only available source of funds needed to complete the Datran system.20

If one assesses the pattern of entry in the post-Carterfone-MCI years, it is evident that after an initial period of fanfare, realized entry has been a faltering and hesitant process. A number of ex planations are possible. First, it may be that the gestation period for successful entry is significantly longer than the five years which have elapsed. Second, the new entrants may have imperfect information about the nature of consumer needs or the technology of supply. Third, it is possible that the established carriers enjoy pervasive economies of scale which foreclose successful entry, and that the

---


---

Table 1: Historical Data Regarding Bell System Provided Terminal Equipment, 1963 – 1973

<table>
<thead>
<tr>
<th>Years</th>
<th>PBXs</th>
<th>Data Base</th>
<th>Type of Equipment</th>
<th>Accounting</th>
<th>Key Indicators</th>
<th>Annual Growth Rate</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1963</td>
<td>1311</td>
<td>20,967</td>
<td>407</td>
<td>23,359</td>
<td>16,820</td>
<td>14.3%</td>
<td>110,36</td>
</tr>
<tr>
<td>1964</td>
<td>1418</td>
<td>22,915</td>
<td>404</td>
<td>24,860</td>
<td>16,820</td>
<td>12.8%</td>
<td>123,67</td>
</tr>
<tr>
<td>1965</td>
<td>1542</td>
<td>25,951</td>
<td>404</td>
<td>27,476</td>
<td>16,820</td>
<td>15.3%</td>
<td>137,69</td>
</tr>
<tr>
<td>1966</td>
<td>1673</td>
<td>29,100</td>
<td>404</td>
<td>29,782</td>
<td>16,820</td>
<td>14.6%</td>
<td>152,26</td>
</tr>
<tr>
<td>1967</td>
<td>1877</td>
<td>31,880</td>
<td>404</td>
<td>32,221</td>
<td>16,820</td>
<td>15.4%</td>
<td>167,52</td>
</tr>
<tr>
<td>1968</td>
<td>2025</td>
<td>35,004</td>
<td>404</td>
<td>35,836</td>
<td>16,820</td>
<td>15.0%</td>
<td>183,30</td>
</tr>
<tr>
<td>1969</td>
<td>2199</td>
<td>38,104</td>
<td>404</td>
<td>38,584</td>
<td>16,820</td>
<td>15.3%</td>
<td>200,45</td>
</tr>
<tr>
<td>1970</td>
<td>2393</td>
<td>41,331</td>
<td>404</td>
<td>41,625</td>
<td>16,820</td>
<td>15.1%</td>
<td>218,76</td>
</tr>
<tr>
<td>1971</td>
<td>2588</td>
<td>44,660</td>
<td>404</td>
<td>44,976</td>
<td>16,820</td>
<td>15.4%</td>
<td>239,05</td>
</tr>
<tr>
<td>1972</td>
<td>2796</td>
<td>48,099</td>
<td>404</td>
<td>49,432</td>
<td>16,820</td>
<td>15.1%</td>
<td>261,25</td>
</tr>
<tr>
<td>1973</td>
<td>2991</td>
<td>51,560</td>
<td>404</td>
<td>52,970</td>
<td>16,820</td>
<td>15.1%</td>
<td>284,80</td>
</tr>
<tr>
<td>1974</td>
<td>3197</td>
<td>55,004</td>
<td>404</td>
<td>56,432</td>
<td>16,820</td>
<td>15.0%</td>
<td>310,52</td>
</tr>
<tr>
<td>1975</td>
<td>3401</td>
<td>58,460</td>
<td>404</td>
<td>59,898</td>
<td>16,820</td>
<td>15.1%</td>
<td>338,07</td>
</tr>
<tr>
<td>1976</td>
<td>3607</td>
<td>61,961</td>
<td>404</td>
<td>63,396</td>
<td>16,820</td>
<td>15.0%</td>
<td>367,03</td>
</tr>
<tr>
<td>1977</td>
<td>3814</td>
<td>65,498</td>
<td>404</td>
<td>66,992</td>
<td>16,820</td>
<td>15.1%</td>
<td>400,42</td>
</tr>
</tbody>
</table>

The literature on economics of scale in communications is fragmentary and far from conclusive. Trowd argued that traditional switching equipment (i.e., crossbar and step-by-step) was subject to increasing costs. C. E. Mayo, "Economics of Public Utility New York: Sinacher & Co., 1947) p. 27. Theo and Lewko found that the cost per subscriber line displays an uncorrelated pattern for exchanges with up to 100,000 subscriber lines. Above that, exchange apparently increased costs. D. A. Bowers and W. F. Lovley, "Diagnosability and Increasing Costs: A Study of Local "Telephone Services," Land Economie 41 (February 1950): 21 - 60. Littlechild argued that there were three sources of economies of scale for local long transmission. These included the use of trunk, the substitution of microwave and coaxial cable for pole systems, and planning and coordination. S. C. Littlechild, "Peak Load Pricing of Telephone Calls," Bell Journal of Economics and Management Science 1 (Autumn 1970): 200 - 203. These arguments have very limited direct relevance to the problem of determining whether specialized carriers and interconnected systems suffer from inherent diseconomies. A definitive analysis of economies and diseconomies of scale in communications still remains be developed.


Harry Trebing and William Melody 105

mine whether there are factors at work, other than scale economies, which will distort the outcome of the FCC entry policies. To explore the role of these factors as determinants of entry, it is necessary to examine the response of the regulated firm to new entry into the market and the remaining barriers to entry in domestic communications.

Entry under Regulation

In a nonregulated sector, the established oligopolistic firm has a number of options when confronted by a potential entrant. The oligopolist may set prices at short-run levels designed to maximize profits and encourage entry, or it may choose to practice some form of limit pricing. If the oligopolist sets prices below the entrants' minimum average cost, then there will be complete foreclosure. The overlap between the price set by the existing firms and the lowest average cost of the entrants will influence the rate of new entry. Conventional wisdom appears to suggest that it is most profitable for the oligopolist to set the price sufficiently above the strict entry-deterring level so as to "encourage a modest amount of entry, but not so much above it as to precipitate a flood." Similarly, the multiple product firm will tend to set prices very close to the short-run profit maximizing level in markets where it has no inherent long-run cost advantage, while it will set prices below profit maximizing levels and close to entry-deterring levels in markets where the established firm has some cost advantage, and it will set prices at profit maximizing levels in markets where the firm enjoys substantial cost advantages. There are four distinctive features of limit-entry pricing in this nonregulated, oligopolistic setting. First, it reflects some loss of short-term profits. Second, it involves a trade-off be between current profits and the prospect of long-term benefits associated with maintaining oligopoly power. Third, there's the prospect that the market share of the dominant firm or established oligopolist group will tend to diminish over time. Fourth, there is the prospect that increased entry in a nonregulated sector will grow to the point where price is depressed toward competitive levels. This will depend, of course, upon the output policies of established firms and the feasibility of...
in the monopoly markets depending upon the rate of growth in demand and the skillful application of second and third degree price discrimination.

Under these circumstances, a simple relaxation of entry restrictions cannot be expected to yield discernible results in the long run. Furthermore, the behavior of the regulated firm vis-à-vis new entrants takes on a significance which is much broader than limit-entry pricing when it is placed within the context of the structural and institutional setting of the communications industry.

Structural and Institutional Barriers to Entry

There are at least six significant entry barriers which confront existing and potential entrants in communications even though the FCC has dropped the legal barriers associated with interconnect and specialized carrier service:

First, as noted above, there is the pricing power of the dominant firm. MCI and other entrants have repeatedly charged that AT&T has employed price as a means of forestalling the growth of competition once regulatory prohibitions are relaxed. Limit-entry pricing holds a clear potential for shaping market structure, and its effective use is enhanced by the fact that a very small portion of Bell revenues comes from markets that could be considered as competitive or susceptible to rivalry.

Bell has moved on a number of fronts to employ price as part of a strategy to meet actual and prospective competition. It has aggressively promoted the inverse elasticity concept together with long-run incremental cost as a rationale for its new pricing proposals. Acceptance would give the carrier virtually all of the flexibility of the traditional value-of-service philosophy with none of the rigidity accountability associated with full-cost pricing. Bell has also introduced a range of post-MCI and post-Carterfone tariffs and pricing concepts. Some have been tried and then withdrawn, but the system-wide averaging and offer high discounts for private line circuits. Others, such as the Hi-Lo Density tariff, mark a radical departure from system-wide averaging and offer high discounts for private line circuits between 370 rate


For a representative example of MCI’s views on AT&T’s pricing practices, see Statement of William C. McGannon, MCI Exhibit 1, In the Matter of American Telephone and Telegraph Company, Long Lines Department Charges, Regulations, Classifications and Procedures for Voice Grade Private Line Service (High Density-Low Density Rate Structure) Filed with Transnational Letter No. 11691, FCC Docket No. 1911.
centers or locations. Still other concepts, such as network access pricing, have been rumored but have yet to appear. At the exchange level, Bell has also vigorously employed price to meet actual and prospective competition. It has required that the customer pay a monthly recurring charge for a protective interface device if the customer furnished its own terminal equipment, but he pays only the tariff lease rate if he uses common carrier terminal equipment (even though both may be virtually identical). Concurrently, the associated companies are gradually increasing the level of charges for the interfaces to reflect the alleged increased cost of labor and capital while reducing their charges for Bell-supplied PBX and key systems and other vertical offerings. The result, as Richard Gabel has shown, is to place the interconnect supplier in a vertical squeeze.

A second significant barrier lies in the inherent structural and operational characteristics of the common carrier communications industry. The need to design and provide service on an interfirm basis requires joint or interdependent planning. Similarly, the need to establish a system for revenue settlements from jointly provided services creates a high degree of interdependence between Bell and the independent telephone companies. Interconnect and specialized carriers are subject to comparable pressures in varying degrees, with the result that a framework of monopoly focal points and tacit oligopolistic coordination is established. In such a setting new entrants can be effectively excluded or admitted under circumstances that severely limit their freedom of action. The need for specialized carriers to work out contracts for local loops and similar interconnection arrangements with existing carriers is a case in point. The specialized carrier is at a disadvantage and vulnerable to delays and other practices which may result in a competitive handicap. To illustrate, AT&T and MCI began negotiations for the provision of interconnection services as a result of the 1971 specialized carrier decision.

Harry Trubing and William Melody

sion. Negotiations continued through the late summer of 1973, at which time AT&T and its affiliates broke off negotiations with MCI and submitted tariffs to each commission in the states where MCI sought interconnection. Pending approval of these tariffs, AT&T announced that it would not provide interconnection to MCI if such services terminated in AT&T supplied equipment.

A third barrier is the impact of increased efforts at product or service differentiation on entry. Individual common carrier offerings have traditionally been viewed as homogeneous service. However, in recent advertising programs, AT&T and the associated companies have sought to differentiate Bell equipment, maintenance, and service from that provided by the interconnects and specialized carriers. The result, as Richard Gabel Bell has sought to create the impression that the entrants' offerings and services are inferior and hence a differentiated product. In this effort, AT&T undeniably enjoys a significant advantage in terms of inherited customer preferences, prior contact, and nationwide coverage. Furthermore, Bell's size gives it a series of special advantages in such nonprice rivalry. Quantity discounts are available from advertising media. Also, successful advertising requires a minimum threshold level of expenditures for coverage and a continuing effort, since the long-term promotional effect is assumed to be cumulative. In these areas, any comparison between AT&T and potential rivals in terms of relative resources borders on the ludicrous.

Fourth, new entrants face the problem of raising capital for expansion. The prospect of high earnings normally appears to be sufficient to attract new funds, and, indeed, the ability of MCI to persevere through more than a decade of litigation and start-up costs is a tribute to this phenomenon. However, there is a real question whether the new entrant can sustain the prospect of such earnings in the face of a concerted campaign of limit-entry pricing and nonprice rivalry by the established firm. Over time, it is reasonable to assume that uncertainty will increase as the new entrant seeks to expand, while AT&T is assured a fair lower cost of capital which avenges the risks of serving a range of monopolistic and competitive markets.

Fifth, the new entrant must continually face the allegation of cream skimming. That is, the entrant will be viewed with suspicion

29 Most of Bell's rate and service changes have been in areas perceived as susceptible to competition. Among others, they include the Hi-Loo Density tariff, Data Under Voice (DUV), and the Digital Data System (DDS). For an extensive evaluation of AT&T pricing practices and philosophy see Robert Triley, "The Economics of Dr. William H. Melody," FCC Docket No. 19120, Phase II, Trial Staff Exhibits.


33As an illustration, MCI has alleged that the Bell Hi-Loo tariff has seriously impaired its ability to raise capital for long-run expansion, while customer diversion has created a cash flow problem for the short-run. See Statement of William C. McGuigan.
as an enterprise only willing to serve large customers in high-density, lucrative markets, thereby burdening the small customers in their markets or possibly jeopardizing the universality of service. The result contributes a strong appeal for denying entry on the grounds of equity and fairness for small users.

Sixth, the new entrant is dependent upon an administrative-judicial process in order to expand service or to protect itself from the retaliatory actions of rivals. This process is slow and costly for a small firm with limited resources, and the outlook is often far from optimistic.65

In addition, the entrant must contend with the continuing confrontation between the state commissions and the FCC over liberalized entry. Many state commissions regard the FCC policy with hostility because it causes exchange and intrastate toll rates to increase, while rates for vertical services fall. This imbalance is aggravated by inflation, and it is not uncommon to find an associated company requesting a massive rate increase for basic services at the very time that vertical services are being drastically reduced.66 A number of states have undertaken full-cost studies to determine the extent of cross-subsidization between these service classes,67 but there can be little doubt that a reduction in charges for vertical services will foreclose the possibility of conscious regulatory cross-subsidization in order to hold down exchange and short-haul toll rates. It is not surprising, therefore, that some states appear to be moving in the direction of banning foreign attachments.68 This prospect can only serve to increase the uncertainty facing the new entrant.

Why Continue Liberalized Entry?

On the basis of historical experience since the Carterfone and MCI decisions, it is not unreasonable to assumethat potential entry may have had more of an impact than actual entry in the interconnect and

---

65The length of time required for commission decision making is a well-established criticism of the regulatory process. What is more distressing is the recent tendency of agencies to defer affirmative action by resorting to a seemingly endless series of contract research studies and interagency commitment to study the problem in further detail. This is particularly true in the case of equipment interconnection.

66Examples are discussed in Globe, "Barriers to Entry." Note especially p. 32.

67The New York, New Jersey, and District of Columbia included their 1972 studies in their 1976 reports on the potential for competition in office communications.

68The New York, New Jersey, and Massachusetts, Vermont, and the District of Columbia have initiated proceedings to develop cost studies for the various vertical services.

North Carolina, Oklahoma, and Utah. See Business Week, 27 July 1974, p. 43.

Harry Trubing and William Melody 111

private line markets. It is also obvious that entry has not been a self-reinforcing force, with the continuing large-scale entry of many new firms serving an ever-expanding series of markets. Yet, the beneficial effects of liberalized entry (whether actual or potential) have been substantial. New services have been introduced, customer options expanded, constraints relaxed, and many traditional practices have been either overturned or reexamined.

An analysis of new technical developments in the interconnect market has been made by Barbara Epstein.69 This study compared the features of PBX equipment manufactured by Western Electric with the features offered on comparable equipment by other manufacturers (largely General Electric, ITT, Norelo, Northern Electric, Ericsson, and Nippon). Epstein found that non-Bell systems offered cost savings to the users in the form of greater convenience, savings in time on the telephone, reduced station equipment requirements, and the need for less floor space.69 She concluded that the relatively poor performance of Western Electric with respect to the features of its PBX equipment stemmed from the slow introduction and marketing of electronic PBX systems. Epstein does note, however, that the Carterfone decision was instrumental in accelerating innovations by Western:

It is only in the past three to four years that Western Electric has been subjected to competition of any dimension. Signs of its earlier growth are already visible. According to a May, 1972, report by the New York State Department of Public Service, "The Bell System also is undertaking the development of new station equipment, particularly paying attention to the PBX area, that will be directly competitive with foreign (non-Bell) manufacturers. To this end it has established a task force, headquartered in Denver, Colorado, devoted exclusively to developing competitive equipment with foreign models both in cost and service features."

A comprehensive statement of the advantages of liberalized entry in the private line market is contained in the FCC notice of proposed rulemaking in Docket No. 18920, Specialized Common Carrier Services.

---


Epstein testimony, pp. 12–18.

ibid., p. 21.
By permitting the entry of specialized carriers, we would provide users with flexibility and a wider range of choices as to how they may best satisfy their expanding and changing requirements for specialized communication service.

There is also a question as to whether the existing carriers can meet the requirements in the specialized markets promptly, efficiently and effectively without prejudice to full and timely satisfaction of the increasing requirements of the public monopoly services. The responsibility for meeting the nation's growing and changing communications requirements is now largely concentrated in the Bell System. This responsibility is becoming more and more difficult to discharge in a manner which enables the Bell System to satisfy timely and effectively all existing and anticipated communications requirements. This is partly because of the diversity of such requirements, the obvious problems of designing and engineering facilities capable of meeting all such requirements with equal efficiency, economy and expedition, and the huge and increasing amounts of new capital the Bell System must raise for construction purposes. The entry of new carriers would have the effect of dispersing somewhat the burdens, risks and initiatives involved in supplying the rapidly growing markets for new and specialized services among a multiplicity of entrepreneurs who appear ready, willing and able to assume these undertakings. It would also expand the capability of the communications industry to respond to the challenges of meeting the rapidly growing and varied demands of communications users.

Further, while economies of scale may result when large general purpose transmission facilities can be used to meet relatively homogeneous communications requirements, there may be other drawbacks. The sheer size of the AT&T organizational structure, its enormous financing requirements, its vertical integration, and new monopoly position in the provision of communications services may make it slower to perceive and respond to individual, specialized requirements and to initiate market and technical innovations. Competition in the specialized communications field would enlarge the equipment market for manufacturers other than Western Electric, and may stimulate technical innovation and the introduction of new techniques. Moreover, new carriers with smaller scale operations could devote their undivided attention to the particular needs to be served and, lacking a captive market, would be under pressure to innovate to produce those types of services which would attract and retain customers.

In an industry of the size and growing complexity of the communications common carrier industry, the entry of new carriers could provide a useful regulatory tool which would assist in achieving the statutory objective of adequate and efficient services at reasonable charges. Competition could afford some standard for comparing the performance of one carrier with another. Moreover, competitive pressure may encourage beneficial changes in AT&T's services and charges in the specialized field, and stimulate counter innovation or the more rapid introduction of new technology.

Of course, it is possible that some forms of entry will yield adverse social costs, but these should only be considered with reference to attendant social benefits. A benefit-cost calculation of this type must precede any reversal of public policy on entry.

Please that regulatory action foreclose entry ignore the fundamental fact that the forces which served to stimulate new entry in the postwar years are still very much at work. These include the rapid growth of demand in existing markets, new technology which increases the demand and new methods of entry, and the lack of demand for a variety of services that have gone unsatisfied. It was the interaction of these forces which created pressures for change in a setting in which the carriers were largely unresponsive. On balance, perhaps the greatest danger in any hardening of entry barriers is that it will preclude potentially efficient and innovative firms from the market.

Major Public Policy Options

What of the future? If no further changes are made in public policy at the federal and state levels, and if AT&T pursues aggressive policies on a variety of fronts, the prognosis for new entry is gloomy.

Existing small entrants in the interconnect and private line markets will probably go through a series of mergers and consolidations (not unlike MCI and N-Triple-C) and face eventual extinction. Potential competitors will not appear after taking cognizance of both pre- and postentry barriers and heavily discounting the prospects for success. The results will appear to support bigness and the preconceptions of those who maintain that economies of scale in all services are universal. It is also unlikely that independent telephone companies will act as aggressive entrants in teleprocessing and private line markets outside of their service areas in such a setting.


**The U.S. Independent Telephone Association has taken a first step in the direction of providing data for an eventual benefit-cost study by establishing an investigation to determine the impact of liberalized entry on 31 independent companies. This study will consider direct and indirect economies of entry by interconnect and specialized carriers. It is too early to discern what data will be generated that will be helpful in evaluating social benefits. See Telecommunications Reports 40 (22 July 1974): 20.
One possible indication of a future trend is contained in IBM’s announcement in mid-1974 that it will form a subsidiary to buy the interests of Lockheed Aircraft Corporation and MCI in a domestic satellite venture which the two companies had previously formed with Comsat. Acquisition of the Lockheed-MCI share by IBM could conceivably place AT&T and IBM on a collision course. Both firms would wish to sustain revenue growth rates of 12 percent per year, and new applications of teleprocessing which combine information processing and communications functions appear to be areas of great mutual interest.

Yet it would hardly seem likely that open warfare would result. AT&T is one of IBM’s largest customers, and AT&T’s chairman, J. D. deBute, is reputed to have said: “Let me say emphatically that we have no problem with IBM’s taking over Lockheed’s and MCI’s share. Their technology is good.” Furthermore, deBute is reported to have discussed the move with IBM’s chairman and indicated that it will “have absolutely no effect on other business relationships between the two companies.”

The prospects are good that a duopoly, based on a policy of market shares, will emerge. Each participant will recognize the other’s sphere of influence and the need to coexist. Whether the resulting duopoly of giants would represent the employment of entry as an effective substitute for regulation can best be left to the reader for speculation.

However, there are other options for public policy with respect to entry. One of the most reasonable alternatives is to recognize that some affirmative action must be taken with respect to the variables that impinge on the entry decision if liberalized entry is to be given an adequate test. Specific attention should be directed to the following points.

1. Establishment of explicit standards of accountability for determining the revenue contribution by class of service for both inter- and intrastate services. This is necessary to minimize cross-subsidization and curb the arbitrary allocation of overhead costs and the distribution of the gains of new technology as well as the burden of obsolescence on the basis of the presence or absence of selective competition.

2. Establishment of rate structures and tariff conditions that negate the capacity to engage in price discrimination. This can be accomplished, in part, by greater reliance on arbitrage and similar practices (as in the case of unlimited Telpak sharing) which erode product differentiation and market segmentation. It can also be achieved by greater attention to the relationship between the pricing of particular products (such as interface devices) and market structure ramifications. On balance, price should gravitate toward cost-of-service conditions and not be a function of discrimination, limit-entry strategies, or an umbrella support program developed by public agencies to preserve a given group of firms or a pattern of market shares.

3. Assurance of equal treatment for all comers in matters of access to and the supply of facilities and services. This includes affirming that common carrier status applies both to service and to the provision of plant and facilities to entrants without restrictive terms or conditions that will tend to make market structures more rigid.

4. Recognition of the interrelationships between the level of earnings and implications for competitive entry. For example, the uncorrected acceptance of simple automatic adjustment clauses which translate all cost increases into rate increases enhances the capacity of the regulated firm to compete with potential entrants who are subject to competitive market constraints.

5. Reconciliation of major differences between federal and state regulatory agencies so that the fears of the state commissions regarding the burdening of basic services will be minimized. This will alloy the prospect that state regulation will be employed as a legal barrier to entry that effectively offsets liberalized entry policies at the federal level.

Conclusion

Although the types of actions noted above should serve to implement the original intent of the Carterfone and MCI decisions, it appears clear that entry will continue to take place at the margin in communications. Significant inroads by new entrants will depend upon the rate of growth in demand, exogenous technological advances, and a major diminution of monopoly power. The latter is, of course, closely associated with the present structure of the industry. In the absence of any significant change in the structure of the domestic communications industry, it is doubtful that entry as a tool

Harry Trebing and William Melody

*See Business Week, 27 July 1974, p. 44.
*IBM currently sells its electronic automatic switchboards (PABXs), telephones, and accessaries in Europe but not in the United States. Market share policies are not new in communications. Between the late 1870s and 1909, AT&T and Western divided the voice and record communication markets by mutual agreement.
of regulation will be strong enough to bring about a dramatic change in market shares or industry organization. The same conclusion seems applicable to any proposal for the general relaxation of entry barriers throughout the industry in an effort to promote expanded entry.

It is also doubtful that commission regulation is currently capable of assuring that potential benefits of liberalized entry in the interconnect and private line markets (whether measured in terms of accelerated innovations or gains in operating efficiency) will flow through in any significant degree to the basic monopoly services. A fundamental revision in the focus and mechanics of regulation will be needed to achieve this objective. For the short term, the best hope would appear to be that of minimizing any burden on the monopoly services stemming from carrier involvement in competitive markets.

In summary, it is reasonable to conclude that the entry problem in communications is essentially a subset of the broader questions of industry structure, corporate conduct, performance, and the appropriate role for social control.

VI. Competition and Entry into Banking Markets

Thomas G. Gies

Following a lapse of almost three decades, there has been a strong renewal of interest in the competitive behavior of commercial banks as well as other financial institutions. The Keynesian revolution focused attention strongly on the macro dimensions of banking, and both theoretical and empirical efforts concentrated on the relationship of monetary institutions to national economic aggregates. Beginning in the early 1960s, however, researchers again directed attention to the micro dimensions of bank behavior. These studies have been oriented in the main toward the structural characteristics of the market as they bear upon competitive behavior—dimensions of the market, number of firms, relative size, economies of scale, and concentration of control. Ease of entry and exit has received somewhat less attention.

Very little effort has been directed toward the effect of nonstructural variables such as management upon the behavior of banks. Economists do not deny the relevance of management attitudes and philosophy in economic performance, but perhaps it is felt that man—
Entry into Banking Markets

Agymion philosophy can only be regulated indirectly through altering market structure and organization. A different argument will be presented here through a case study of a banking group, the COMAC Group, which introduced a highly aggressive type of management in certain Michigan banking markets during the 1960s. First, the actions taken by this owner-management group are spelled out, and the results for the banks under their control are identified. Second, an attempt is made to sort out the impact which these actions produced in three Michigan banking markets—Lansing, Macomb, and Kalamazoo—where the group under study operated. Third, the reactions of the banks involved and, finally, the response of government regulatory authorities to the disequilibrating group will be examined. The lessons for competition and regulation have some force for other banking markets.

Toward a Banking Market Definition

It has long been recognized that banking has oligopolistic tendencies, but only recently has work been directed toward defining a banking market. A group of studies [1, 3, 7, 8, 9, 12, 13, 14] shows that banking is local in nature, with only the largest firms having banking mobility. Most customer affiliation with banks tends to be based on convenience and quality rather than price factors. Large businesses, on the other hand, are more mobile in their bank selection and the primary determinants of their choice are the financial condition of the bank, the location, the number and quality of services offered, and credit availability. It is interesting to note that even in this market segment, which is considered to be the most price elastic, price factors are not of primary importance. This is not to say that large businesses are indifferent to price differences, but they also consider nonprice factors in choosing their banks.

The product markets in which price factors are important are those in which bank customers have nonbank as well as bank alternatives. For example, savings and loan associations compete for time deposits, the money market competes for commercial deposit money, and nonfinancial corporations compete for government securities [5]. On the other hand, such products as demand deposits and business loans come under little nonbank financial intermediary pressure. Competition for these products sometimes takes a nonprice format, with emphasis being placed on new services, product differentiation, and mass promotion.

Thomas Gies

Many studies have been conducted with the aim of determining the effect of bank structure on performance and assessing the effects of new unit banks, de novo branches, mergers, and holding company acquisitions. While there are numerous differences among them, the four discussed below illustrate the principal directions which these analyses have taken. These efforts are of special interest because each touches on problems related to the banking markets under review here. The four studies were conducted by Paul Jessup, Robert Lawrence, and Samuel Talley. Jessup [6] has analyzed the performance of a group of banks where significant shift in ownership has occurred to determine whether operating results were measurably altered by the restructuring of ownership. Lawrence [10, 11] conducted two studies on bank holding companies aimed at identifying differences in performance between banks affiliated with bank holding companies and comparable independent banks. The last of the four precedent studies was conducted by Talley [15] and used Lawrence’s methodology. Talley’s study shows that the major effect of holding companies on operating performance is to alter the portfolio composition of acquired banks.

Price and Output Determination

The Banking System

It is essential to recognize that output and price determination in commercial banking markets bear less relationship to the number of banking firms in the market than is the case with industrial firms. The reason for this springs from the fact that increasing the number of banks has no direct or immediate effect on the capacity of the banking industry as a whole to create credit, for the monetary authority has, at least in the short run, effective control over the total quantity of money and bank credit. Because credit in the United States is the most important product banks produce in terms of revenue dollars generated, the insensitivity of output to number of firms means that the price of bank output also is not directly and immediately related to the number of banking firms in the industry. This single fact is of paramount importance in judging questions of organization and structure, of entry and exit, and of stability in ownership and management in banking markets.

*Whether the number of firms in the banking industry per se affects the attractiveness of bank credit vis-à-vis nonbank credit will not be debated here, although the likelihood seems remote to this author.
The underlying reason for the irrelevance of the number of firms to output of bank credit rests on two facts. First, government policy, as administered through the monetary authority, establishes maximum output of bank credit for the entire economy, although not for individual market areas or for individual banks. Second, the output quota established by the monetary authority is virtually always below the profit maximizing equilibrium for the banking industry, and actual output therefore normally approximates the maximum permitted output. The phenomenon of full utilization of bank reserve positions is the normal condition of the banks and follows from the profit maximization incentive driving the banks to strain against the upper limit on output presented by monetary controls.

The rationale for monetary control is the need to regulate the overall level of employment, output, income, price level, and growth in the economy. The firms composing the banking sector merely represent the vehicle for execution of monetary policy, and the number of firms is irrelevant to accomplishing monetary goals. The addition or subtraction of one firm will not alter central bank decisions regarding the optimum quantity of bank credit at a given time or the optimum growth rate of that credit.

If the above argument is accepted, that is, the total supply of bank credit is outside the discretion of the commercial banking industry, and no decision by the industry can alter maximum output, attention can be directed to the demand for bank credit. In an aggregate sense, this is a function of demographic and economic factors, such as number of households, businesses, and governments and the scale and timing of their financial deficits. Clearly, these demand considerations are also outside the discretion of the banking industry.

If neither the level of nor changes in supply and demand are controllable by banks, it must follow that banks have little potential for controlling the yields on securities traded in open markets, for example, the yield on Treasury issues. Rather, the price of securities is a decision made at the level of the monetary authority and is one of the principal, "proximate" objectives of monetary policy; the ultimate objective, of course, being price and output behavior in goods and service markets.

The whole question of competitive relationships in such a market, the significance of the number of competitors, and their stability, duration of life, and management policies takes on a different light with this in view. The first rule of monopolistic behavior, that output should be reduced from the competitive market equilibrium, may require modification in a market in which maximum permissible output is already decreased substantially below the profit maximizing level even for a monopoly firm. The modern central bank has exclusive control over creation of legal reserves for the banks and by this means can adjust the quantity of bank credit up or down. The presence or absence of vigorous competition in banking markets would not necessarily lead to any difference in output, where output of the industry is severely inhibited by monetary control. Specifically, if the quantity of bank reserves limits creation of bank credit to an amount less than the profit maximizing level, then, regardless of market organization, there will be no further overall restriction of output at the commercial bank level, even where there is only one commercial bank in the market.

Regional and Local Banking Markets

Determination of output and pricing policy within regional and local banking markets arises from considerations at the system level. The monetary authority is responsible for establishing the total quantity and rate of growth of bank credit, but does not specify the proportion of that total which will become available to a regional or local market. However, as with the system of banks, that determination is not a function of the number of banks in the regional or local market in the sense that adding a bank does not add to the supply of bank credit in a given geographic or functional market. This is not to say that structural factors such as the number of bank offices in the market and the relative yields on substitute assets do not affect demand for deposits. In addition, income, wealth, amount of business located in the market, and proximity to other financial centers influence the volume of deposit funds.

But before discussing these forces, it may be helpful to clarify the concept of a banking market. There are two dimensions: geography and product. The geographic market is generally local in nature, although for some customers (for example, very large businesses) the market may be national or international. A considerable volume of studies testifies to this fact. There is much less agreement on what the appropriate product market is. Briefly, the two alternative viewpoints are the single product and the multiproduct concepts. The former, also known as the customer relationship concept, emphasizes

---

3The modern day central bank's control over supply and price of bank reserves is, in fact, one of the few cases of virtually perfect monopoly.
the tying nature of bank products. The commercial bank is seen as providing a bundle of services and consequently is not in direct competition with other financial institutions. Proponents of this view feel that the central service is providing a means of payment, with the other services tied to this essential service. The multiproduct or institutional investor concept sees the commercial bank as providing different products, each distinct from the others, for example, mortgage credit, installment credit, and so forth. In each product area the bank may encounter competition from other financial institutions, and in this respect commercial banks are seen as not being unique.

Actually, the controversy appears to be more one of degree than kind. David Alhadef [26] and Clifton Kreps [9], both proponents of the multiproduct approach, agree that business customers of the bank are often required to keep compensating balances, which is a method of tying products. Even though some of the other bank customers are not required to do all their business with one institution, this might still be their normal behavior. Certainly banks prefer to lend to customers they already know. On the other hand, some credit products such as residential mortgages are quite competitive between banks and nonbanks. Thus, elements of both theories seem to have validity.

Our purpose, however, is not to evaluate the literature on bank markets, but merely to point out the relevant dimensions in identifying the markets in the three study areas under review. Geographically, the three Michigan cities of Lansing, Kalamazoo, and Muskegon are distinguishable, isolated markets. The Kalamazoo Standard Metropolitan Statistical Area (SMSA) makes up one county (Kalamazoo County); approximately 48 percent of the county population is within city limits. The Muskegon SMSA also makes up only one county (Muskegon); approximately 31 percent of the county population is within the city limits. The Lansing SMSA is comprised of three counties (Clinton, Ingham, and Eaton), about 36 percent of the SMSA population is within the Lansing city limits, and 71 percent is within Ingham County. Because of the nature of this study, the product delineation appears to be of subordinate importance.

Given the funds in the market, and thus the potential size of the local banking sector, total credit will be a function of the amount of reserves needed to cover deposit liabilities and the amount of excess reserves the bank desires to keep. The composition of this credit will be a function of such variables as capital requirements, relative

Thomas Gies

yields, perceived risks, liquidity, deposit variability, management preferences, legal restrictions, market quality of assets available, and tax considerations. As Donald Hodgman [16] has pointed out, the customer relationship is extremely important in its effects on loan composition since the bank is likely to favor lending to its major depositors. Looking at the variables that affect output, one can see that individual bank management can control to a large extent both the total quantity of credit extended and the composition of this credit.

Granting that individual banks have substantial control over output decisions, how much control do they have over pricing? Without becoming too specific, certain generalizations can be made concerning the pricing policies for different credit outputs. Except for small, locally placed, municipal issues, the individual bank has no control over pricing for marketable securities. These rates are determined on the open market and are exogenous to the individual bank.

On the other hand, commercial banks in general have power to set prices on certain assets, such as customer loans. For large business loans, the market is demonstrably national in scope, which means that besides the large number of banks to choose from, national firms also have the money market as a viable alternative to short-term borrowings from commercial banks. For small businesses, however, credit sources are much more restricted geographically, and the local commercial bank has more autonomy in setting rates.

The market for mortgage credit is divided between commercial and residential real estate. Commercial real estate loans often draw lenders nationally, as for example, a large regional shopping center, while single family mortgage credit is largely restricted to local markets. Residential mortgages are relatively homogeneous, prices are typically close for all suppliers of this type of credit, and consequently, banks do not have much freedom to set rates. The increased willingness of out-of-town nonbank mortgage lenders, such as life insurance companies, to lend in a national market through local mortgage brokers has further unified mortgage rates.

Consumer installment credit demand is regarded as inelastic to price in the aggregate, but this may be less true with respect to the individual bank. Some borrowers will go to one bank and accept whatever terms are offered, while others shop around. They may be

*Although they have more power to do so, it is not clear whether they exercise this power. Studies on the relationship between concentration and loan rates are inconclusive.
willing to pay 20 percent per year for credit, but they are not insensi-
tive to price differences. Recognizing this, banks appear to discrimi-
nate among classes of borrowers on the basis of competitive condi-
tions rather than on cost alone. Summarizing, it can be said that while total deposits in the mar-
ket are determined mainly by factors outside the banks’ control, an effective liability management program may enable the bank to increase its supply of funds and thus its credit output. Furthermore, by increasing its willingness to take risk, the bank can further minimize its liquid reserves and increase credit output. In addition to increasing total credit output, the bank can change the composition of this output. By shifting from low yield credits to higher return uses, the bank may be able favorably to affect its revenues. The important point is that both the increase of total output and the change in its composition are within the scope of management powers of the indi-
vidual commercial bank.

Static versus Dynamic Market Organization for Banks

The foregoing considerations form the basis for arguing that the structure of the banking market may have less influence on the price and output of bank credit than is the case for industrial products. Rather than sheer number of competitors, the important dimension of bank markets may be (1) ease of entry and exit and (2) managerial philosophy. The magic quality needed to assure vigor in exploring new products and new methods may be more intimately related to change in the number of competitors than to the particular number, whether that number is large or small. In other words, fewness of banks may be less disadvantageous than is generally thought, pro-
viding the membership of the small group is capable of being changed readily or has a genuine potential for change. Beyond this, the man-
gagement philosophy of the small owner-manager group may be a critical behavior dimension that is not revealed in standard market structure models.

It is a common observation that local retail merchants become accustomed to each other’s presence — their respective polices con-
cerning pricing, quality and service, and so forth — and feel little threat from the rival’s operation. But let word get around among pharmacists, for example, that a new drug store is going to open, and deep concern is aroused: What will the new store offer? Will it cut prescription prices? Open on Sunday? Where will it locate? and so forth. The addition of a new firm is disturbing, whether the existing number of firms is one, two, or ten.

The principal direction of effort to establish control in a market may not be to wipe out existing competitors, but to prevent new competitors from entering. Existing competitors are known quan-
tities and in many cases have come to accept the very human ap-
proach to life of “live and let live.” But since no one knows exactly what a new competitor will do, the conventional oligopoly wisdom would be to control entry.

The banking business achieves such controlled entry through legally defined and sustained barriers. For reasons long established and well known, entry to the banking business requires obtaining a license from the state. In most jurisdictions, existing banks are pro-
vided extensive opportunity for objecting to new applications for entry. The existing firm or firms can be counted on to allege that they are “adequately serving” the market and that there is “no necessity” for additional facilities. In many jurisdictions, “necessity” is interpreted as meaning that no existing bank will be injured, and in one recent instance, the issue involved the question of whether the new entrant might injure future growth potential of existing firms (banks).

Empirical Studies of Bank Markets

A review of some of the empirical work on bank market entry provides insight into why market entry occurs and the impact of new

competitors. Before examining the research noted earlier, two studies which have examined the reasons for entry into a particular banking market deserve attention.

Robert Chandraos [24] looked at the performance in certain bank-
ing markets before and after new banks entered these markets. The performance ratios of these entry markets were compared to an average ratio composed of all nonmember banks in that particular state. Findings revealed that earnings were higher in the entry than in the nonentry markets for the period preceding entry. Furthermore, loans-to-assets ratios were lower in the entry markets. It might be hypothesized that the reasons for entry were wound: to take advan-
tage of the already greater than average profits in the market and to

further increase these profits by increasing the below average loans-
to-assets ratio.

In the other study, by Donald Fraser and Peter Rose [4], entry and nonentry markets were compared in order to determine if there were
different performance characteristics. The intent was to find out why new banks wanted to enter these markets. The Fraser and Rose results conflict in some respects with those of Chandross. Although the loans-to-assets ratio was found to be lower in the entry markets, there was no difference in profitability between the entry and nonentry markets. In addition, Fraser and Rose found that the entry markets had a higher ratio of U.S. government securities to total assets and a lower ratio of time deposits to total deposits.

Neither Chandross nor Fraser and Rose looked at the situation of entry through the holding company or chain banking mechanism, but we feel that some generalizations may be made concerning bank entry motivation. From both of these studies it can be seen that banks entered the markets either because of the high level of profits or the potential for profits that could be realized through new management policies (for example, portfolio shifts). Fraser and Rose conclude that "entry of new banks into Southwestern communities did not appear to result from excessive profits carried by the previously operating institutions. Instead, it appeared to stem from the anticipation of the probable streams of profits arising from an expansion of credit in the local area" [4].

Numerous studies have attempted to determine the impact on bank performance of a change in market structure. Of interest here are the four studies noted earlier, three of which looked at the impact of holding company acquisition, and one of which examined the impact of changes in bank ownership and management. Jessup [6] analyzed banks which had a new majority ownership and new senior management and compared them to control groups of other banks both before and after the changes. Jessup hypothesized that important differences in performance are associated with new individual ownership of banks, and he was able to confirm this by statistical tests. In general, the banks studied experienced higher loans-to-assets ratios, a higher percentage of consumer loans, and no change in the prices of bank services. Thus it appears that the supply of credit to the local economy was increased without additional cost to customers.

Lawrence [10, 11] conducted two studies concerning the effects of bank holding company acquisitions on performance. The first used the paired method of analysis. Banks that had been acquired by holding companies were compared to similar, nonaffiliated commercial banks both before and after the acquisition. The results indicated that differences in performance were minimal. Holding company affiliates did, however, have a higher ratio of municipals than nonaffiliated banks. Furthermore, there was a smaller percentage of assets committed to U.S. government securities and to balances due from domestic banks. The breakdown for banks shows that installment loans increased more than any other category. From these and other statistical results, Lawrence concludes that bank holding company affiliates tend to be more aggressive lenders than independent banks and thus make more credit available to the local community. It is interesting to note that there was no significant difference between affiliates and nonaffiliates on interest paid to savings depositors or interest charged on loans, although affiliates charged higher interest rates on demand deposits. Finally, profitability differences were found to be insignificant.

Lawrence's second study looked at the operating policies of the multibank holding companies toward their banks. To accomplish this, an expensive questionnaire was distributed to all registered bank holding companies. Lawrence found that they were, in fact, important differences in the operating policies of these firms, the degree of centralization of decision making varied among holding companies, and these differences could not be attributed to any economic or structural factors. Instead, Lawrence suggests, "a holding company's policies can only be determined by investigating the management philosophy of the senior officers of the particular company" [11, p. 32]. There were, however, some areas where holding companies exerted more influence, including securities investments, federal funds transactions, and bank correspondent relationships. Other functions, such as pricing policies, decisions on the composition of the loan portfolio, and decisions on individual loans, were influenced least by holding companies.

Finally, Talley [15] updated the first Lawrence study on holding companies, which had used the paired method of analysis. Talley found that affiliates had significant differences from independent banks. After acquisition, for example, affiliates reduced their ratio of cash to total assets and their ratio of U.S. government securities to total assets. More funds were shifted into state and local government securities and loans, particularly installment loans. Consequently, it was hypothesized that affiliates better served the credit needs of the local community. Little change was found in the amount of capital provided and the profitability of the affiliates. Although Talley found affiliates' service charges on demand deposits to be lower, the result was not significant at the .05 level. This conflicts with Lawrence's
Finding that service charges on demand deposits rose. As in the Lawrence study, prices for savings accounts and loans did not change significantly. Thus, except for the service charges, the results of the Talley and the Lawrence studies are similar.

We can say that, in general, we would expect acquired banks to reduce their liquidity as measured by cash and U.S. government securities and increase their investment in state and local securities and loans. We would expect to find little change in capital, prices, and profitability.

Criticism of Contemporary Bank Management

The U.S. banking system has been changed significantly in the past 25 years by a number of important factors, such as the growth of nonbank competitors and, indirectly, changes in the U.S. economy. On the other hand, a number of observers, among them George Mitchell, Member of the Board of Governors, Federal Reserve System, have noted that the banking industry itself has not been an innovative force for change. Mitchell has written that some firms and industries have produced a product or service, generated public awareness and acceptance for it, and using generative, adaptive, and creative forces from within, have accomplished a role and importance for their own enterprise. The result is an identifiable new business or industry. By contrast, in the case of banking, Mitchell feels that there has been an absence of innovative quality:

"Banking is not such an enterprise or industry. It has had a pattern of traditional services, an imposed molecular structure, and a pedestrian operating technology, none of which it could call its own. It has not innovated its service products or shown much adaptive ingenuity in their promotion. Its favorite image has been a passive conformity to the moves of its better customers. Its competitive aggressiveness has been schizophrenic, with large sectors of the industry advocating or supporting publicly administered price ceilings for time deposits, public prohibitions against the absorption of exchange, and a variety of regulatory devices or postures that by sanction or promise dilute competitive ingenuity."

Another author has described bankers as follows: "Aggressiveness and profit motivation have become synonymous with recklessness; backwash unwittingly became synonymous to conservatism and prudence. The management of a commercial bank is not on the whole very difficult. It calls for prudence, probity, adherence to routine and system, and large acquaintance in the business community" [18, pp. 62–63].

While considerations other than bank management inadequacies share responsibility for the relatively slow growth of banks, in the opinion of this writer, the fact remains that commercial banks have declined in relative importance and that studies of bank attitudes at about the time the COMAC group was formed tend to bear out the picture of banks as being conservative and traditional in their approach to the provision of credit and other financial services commonly used by the U.S. economy. Federal Reserve studies of growth of financial intermediaries reported in the Flow of Funds Accounts indicate that commercial banks held approximately 70 percent of the assets of all major financial intermediaries in 1945, but 25 years later, in 1970, that share had declined by 54 percent.

L. F. Whitledge [22] surveyed commercial banks in the early 1960s with a view to identifying bank attitudes toward marketing against competitive institutions and other measures of management aggressiveness. Of interest is his finding that 31 percent of the banks felt that encouragement of mass use of installment credit by individuals for purchase of consumer goods was inconsistent with encouragement of personal integrity and good character. "A bank which has an aggressive consumer credit installment loan policy wholly for the purpose of making profit for the bank is not violating any duty to the public, the banking industry, ownership or customers." A full 40 percent of bankers surveyed felt that this statement was not valid. Nearly 25 percent of bankers surveyed indicated that consumer credit was not consistent with good banking, bank safety, and the public service required of a bank.

The reported views refer only to the Whitledge sample, of course, and may not be fully representative of the universe of banks. Nevertheless, it seems ironic that as recently as the last decade, a sizeable portion of any sample of banks considered consumer lending to be a questionable form of credit. After all, a bank's business is lending money, and the fact of the matter is that the household sector of the U.S. economy has grown vasty in its need for credit as a result of a number of influences, including the extraordinary quantity of funds needed for the construction of new housing and for the financing of new cars and durable household goods. In fact, the enormous increase in credit demand from the household sector may be the most striking feature of the U.S. financial scene during the past 25 years, yet the banking industry was slow to recognize the opportunity and lost the
momentum of growth to the savings and loan industry, the sales finance companies, and the credit unions.

From another area of banking surveyed by Whittington it is interesting to note the results of a question that relates directly to the matter of competitive attitudes of bank management. Banks were asked the following about new business solicitation of commercial customers: "Is there a general policy against raiding competitor bank customers?" Fully 26 percent of the group stated that their firm avoided raiding other banks' customers; an additional 7 percent replied that they "didn't know." Thus, more than one-third of the group either had a definite policy against aggressive competition, or no policy at all (at least not a known policy). This fact seems especially significant, for commercial customers represent the market in which banks have been the specialists. Certainly one could not have called the above attitudes aggressive.

It seems a fair assumption that the decline in relative importance of commercial banks over the last 25 years can be only partially explained by the changing economy. It also should be recognized that banks have failed to respond fully to the changes that have occurred in the economy and thus have made the decline greater than it might otherwise have been. Banks have aggressively pursued new business only when someone else has demonstrated the value of such business. David Rockefeller, Chairman of the Board, Chase Manhattan Bank, has noted this characteristic: "But as the years went by and the demand became increasingly complex, many banks did not respond as alertly as they might have. They held back from new fields, turned away from new services. Inevitably, therefore, specialized institutions such as credit unions and savings and loan associations moved in to fill the vacuum. Subsequently they also became major competitors of the commercial bank."

**COMAC: Origin and Functions**

COMAC, an acronym for Comprehensive Management Company, was formally brought into existence in 1967 for the purpose of systematizing the management functions in a rapidly growing conglomerate engaged mainly in banking and real estate. In an informal sense, the group of partners which formed COMAC had been in operation for some five or six years previously, but the need for formalizing the group arose from the complex planning, coordination, and direction required for the successful operation of a group of ten banks, a number of large commercial real estate properties, and, eventually, three overseas investment offices. The functional relationships within the group are illustrated in Figure 1.

The COMAC partners were interested in acquiring investments in which they would have a controlling interest and in which they would have active management participation. The focus of their interest was for the most part lay in the financial area, with some interests in real estate markets. As time proceeded, the domain of their operations spread so widely and included such vast numbers of individual business firms, both bank and nonbank, that it became awkward if not impossible for the informal group of a half-dozen to control, directly and personally, the individual firms that composed this federation. When COMAC was formed in 1967 the employees who, together with the original six partners, did the active managing and directing of the properties numbered less than one dozen. The em-
severe competition in the U.S. economy appeared then, as now, to be in the areas of heavy manufacturing and high technology firms. The U.S. manufacturing sector has been a high growth area over the past 50 years, while the high science firms represented the most dramatic and potentially profitable section of U.S. industry during the decade of the 1960s. Nevertheless, COMAC was determined to shun these areas as objects of investment because the reason was simple. They were attractive to too many highly skilled, competitive, owner-management organizations that competition in such industries was likely to be intense. Instead, COMAC elected to specialize in those areas of the economy where growth potential was substantial, but which were receiving relatively little attention from such aggressive firms as Goodyear and Western, Xerox, IBM, Dow, and McDonnell. Second, it was necessary that attention be directed into those areas where it was financially possible to develop an acquisition plan. For example, the mutual organization of the savings and loan and insurance industries made it difficult if not impossible to develop a system for taking control of these institutions. The logical institution was one in which there was stock ownership, but in which that ownership was sufficiently concentrated to permit effective control to be obtained through a stock offer plan. Commercial banks, particularly those of modest size, represented the ideal object of acquisition. In addition, it was quite evident even in the late 1960s that banks were not developing loan, investment, and deposit volume up to the potential that was present in many markets. This was evident not only in country bank markets, but also in metropolitan areas where controlled-entry assured existing firms or not spectacular results.
management group was not bringing the operation up to the potential that the market afforded. Identification of these circumstances could be made from a review of deposit volume, the growth rate of deposits, the loan-to-deposit ratio, the bank's tax position, and the operating costs of the bank. In short, profit position was a revealing measure of the effectiveness of management and ownership in achieving the potential that the market afforded. The ideal situation was one in which there were three or four banks in a market, one of which, frequently a long-time member of the market group, lagged in terms of size, growth, service to customers, and profitability to owners. Such a firm might be available at a bargain price. (A bargain price meaning 125-150 percent of book value of the stock.) In many instances, the existing management-owner group was following traditional practices in pricing, marketing (that is, none), and cost control (that is, none), the result of which was low profits on equity capital, frequently 5-6 percent, or about the same as on a good grade of corporate bonds.

Specific COMAC Markets:
Lansing, Kalamazoo, Muskegon

The COMAC operation eventually penetrated twelve banking markets and directed the activities of 19 banks, including two overseas. However, for purposes of this analysis, the three market areas of Lansing, Kalamazoo, and Muskegon have been identified as exemplifying the COMAC operation. These three were because they are sufficiently distant from Detroit to avoid strong influence from its banks, and yet represent markets of sufficient size and complexity to enable us to analyze the impact of COMAC management efforts. The behavior of these three banks and market areas in terms of performance, growth, development, and change under COMAC direction will be compared to that of rival banks in other communities during 1966-1970.

We might ask the question: If the three banks had continued under their former ownership and management, would they have followed approximately the trend of development of all banks in the state of Michigan? Using the performance of all Michigan banks as a proxy for the independent growth and development path which these three banks would have taken, we can make some specific comparisons.

All Michigan banks during the 1966-1970 period increased in total holdings by some 52 percent. Asset allocations also changed significantly for all Michigan banks during this period: Government securities declined from 23 percent of total assets in 1966 to 13.6 percent by 1970. Obligations of states and political subdivisions, that is, tax exempt municipals, increased from 11.7 percent in 1966 to 12.6 percent in 1970. Loans as a percentage of assets increased from 53 percent at all Michigan banks in 1966 to nearly 59 percent in 1970. Cash assets declined very slightly, from 10.6 percent to 9.9 percent. The ratio of time deposits to total deposits increased from 65 percent to 67 percent in 1970. The capital ratio of all banks in the state increased very slightly from 7.8 percent of total assets to about 8.1 percent. Income after taxes as a ratio to capital accounts was 8.9 percent in 1966, and this ratio increased to 10.25 percent in 1970.

These changes in the distribution of assets and in the sources of funds and operating returns parallel the performance of all banks in the United States. The banking industry as a whole increased in size, shifted asset allocations in the direction of more loans and smaller amounts of Treasury securities and cash, and increased its return on equity capital by modest amounts.

Bank of Lansing

Comparing the changes in all Michigan banks with those at the COMAC banks reveals very sharp distinctions, and a summary of these is revealing. The Bank of Lansing, for example, increased its total assets from the beginning of 1966 to the end of 1969 by approximately 150 percent, that is, from $66 million to $154 million. This surprising change was funded largely from increases in time and savings deposits. These deposits totaled $33.7 million in 1966 and increased to more than $70 million by the end of 1969. In the same period, deposits of states and political subdivisions increased from about $2.5 million to over $16 million, and demand deposits of individuals, partnerships, and corporations increased from $23 million to $35 million. In addition to these deposits sources, the Bank of Lansing employed external markets for sourcing asset acquisitions. It is apparent from condition reports that the bank was heavily involved in federal funds purchases ($6.75 million) at year end 1969. At that time the bank also had developed Eurodollar lending sources. Initially, capital accounts provided $4 million at the beginning of 1966, and $7.6 million at year end 1969.

In addition to these increased sources of funds, asset allocations were shifted drastically during the three-year period under review.
the management of the bank prior to the COMAC purchase was gradually increasing holdings of tax exempt securities as a proportion of total assets, this process was greatly accelerated following acquisition. By the end of 1967 tax exempt municipals comprised more than 17 percent of the assets of the bank, and by the end of 1969 this ratio had risen to 31 percent. As a consequence of these adjustments in cash and security holdings, there was an overall reduction from the early to late 1960s from about 56 percent to approximately 40 percent. However, the change in the composition of the cash and government securities holdings was equally marked. The portion that had been composed of cash and Treasury securities was largely eliminated. In 1965, such assets accounted for 80 percent of the bank's security holdings, whereas by 1969 they represented only 15 percent. State and local government securities had replaced Treasury issues and represented 75 to 80 percent of total security holdings. As noted above, the loan volume at Muskegon Bank and Trust was slightly above 40 percent of total assets in 1961. This ratio also was changed following the acquisition by the COMAC group. There had been a tendency for the loan ratio to rise during the 1960s, and it can be noted that, in contrast to the ratio in 1961, loans had risen to 43 percent by 1963 and to 45 percent by 1965. However, under COMAC ownership and management the loan ratio moved upward even further, and by the end of 1967 it comprised 49 percent of total assets and exceeded 50 percent on each of the call dates following that; the maximum ratio was 57 percent in mid-1968. It is interesting to note that the sum of cash, securities, and loans on the June 1970 call date amounted to 97 percent of total deposits of the bank, which probably was matched at certain other major metropolitan banks, but which certainly was unusual among small and medium-sized banks in the Midwest.

**Industrial State Bank and Trust:**

**The Kalamazoo Market**

Industrial State Bank and Trust was acquired by the COMAC group (actually, of course, the acquisition was made by Kalamazoo Investment Company) in 1964, and the standard reconstitutions of the bank's asset structure was initiated. The municipal portfolio was sharply increased, cash assets and U.S. securities were reduced substantially, and a program of time deposit expansion was undertaken. The time deposit program was based upon the growing awareness among savers that interest on bank deposits was worth having and
that the savings and loan industry was succeeding on a largely
price-based competition. Industrial State adopted a program of solic-
titing both passbook savings accounts and certificate of deposit ac-
counts, generally setting the pace in the Kalamazoo market, for as rates were concerned. A major part of the marketing effort was
aimed at individuals, but business and government deposits were not
shunned. The results were striking. Private time deposits, mainly
from individuals, increased from $20 million to $50 million between
1964 and 1970, while public time deposits increased from less than $5
million to more than $27 million.

Efforts were made to increase demand deposits also and met with
a certain amount of success. During the period 1964–1970, demand
deposits grew 50 percent. However, these deposits cannot be at-
tracted on a straight price basis, as is the case with time deposits
within the limits of Regulation Q. Price bidding is evidenced indi-
rectly by the introduction of low, minimum balance checking ac-
count plans.

The overall COMAC plan for developing a bank’s potential mar-
et was to seek out all sources of funds available, providing from the
least (demand deposits) to the most (equity capital) costly. The pur-
pose was not to minimize the average cost of money, the presumed
aim of the industrial financial manager, but to continue acquiring
funds as long as the cost was below the net yield on marginal assets.
Following this philosophy meant developing sequentially the several
sectors of the banks’ funds market. Demand deposits are the lowest
cost source of funds, followed by passbook savings. Certificate of deposit
funds were the third segment of the deposit market developed and
resulted in substantial enlargement of public deposits, as noted
above. (The ratio of public deposits to total deposits at ISB increased
from 13 to 18 percent between 1964 and 1970.)

In addition to solicitation of deposit funds, the COMAC banks
purchased (borrowed) funds in the federal funds and, when possible,
Eurolollar markets. Purchases in the former market were never in
evidence at Industrial State prior to the change in ownership
management in 1964, but by 1966 were regularly in evidence. The
amounts varied, and call date figures probably are not representa-
tive, but it is clear that as much as 9 percent of total funds in use at
the bank were purchased from the federal funds market at certain times.
Eurolollar borrowings are not identified as such in the bank’s
periodic financial statement, but presumably they appear under the
heading “other liabilities.” Amounts up to $5 million were success-
fully borrowed in that market.

On the asset management side of the balance sheet, the
Kalamazoo bank showed the results of COMAC planning promptly
after the acquisition was completed. The first change was a sharp
build-up in the bank’s loan portfolio, as loan “packages” were sold to
the bank by other COMAC entities. At year end 1964, Industrial
State’s call statement showed a very conservative 41 percent ratio of
loans to total assets (at a time when Michigan banks were around 46
percent and Detroit banks around 50 percent); six months later, the
ratio had jumped to 65 percent and continued at between 60 and 90
percent through 1970.

The ratio of cash assets and U.S. securities to total assets also was
sharply decreased. Cash and governments constituted 43 percent of
assets in mid-1964 and 48 percent at year end. Six months later this
ratio had declined to 29 percent, and one year later to 21 percent.
After mid-1966, a build-up in municipal holdings became evident,
and at year end these represented 19 percent of Industrial State’s
assets, compared with 4–9 percent in earlier periods.

COMAC and the Regulatory Authorities

The regulatory posture toward the COMAC banks changed
gradually from uncertainty and “wait and see” in 1964 and 1965 to
warm support at the time that the Bank of the Commonwealth took
over the defunct Public Bank assets and liabilities in 1966. Later,
however, regulators again became doubtful and, eventually, very
hostile. The latter change occurred partly because of tangible, objec-
tive differences between regulators and the COMAC group concern-
ning proper management. The conflict stemmed partially from the
disparity between a new, different attitude toward asset-liability
management and more traditional views, but the existence of a more
and cavalier manner adopted by COMAC personnel cannot be over-
looked. This writer feels that in no small degree the difficulties which
the COMAC group eventually faced by 1969–1970 were the result of
failure to maintain good relationships, full communication, and sym-
pathetic support from the regulatory agencies. There probably is no
single act which triggered the eventual negative attitude of regu-
lators; rather, their disenchantment could be traced to a variety of
individual management decisions on the part of COMAC.

During 1968, and more intensively during 1969, regulatory au-
thorities sought to change the manner of operation of the COMAC
banks. First, and most important, regulators attempted to limit fur-
Thomas Goss

Even more sensitive, however, were banks which had been lending funds to the COMAC group via the federal funds market; they clearly were made very uneasy by the Federal Reserve statement. While one cannot say that the Federal Reserve was responsible for the ultimate demise of the COMAC operation and for the dissolution of the COMAC banking group, it does seem that there is a link between the Federal Reserve’s statement in early 1970 and the problems which ultimately became catastrophic in magnitude.

The central characteristic of the COMAC banking operation is the fact that it was designed around the holding company plan without employing the legal structure of the holding company. At the time the group was formed, holding companies were not provided for under Michigan corporate law; hence it would not have been permissible to form a corporate holding company. Instead, COMAC represented a management partnership which performed a supervisory, shepherding function over some 19 banks. Each bank was owned primarily by an independent partnership which individuals were identical to or had a substantial connection with the COMAC partners. Hence, the total apparatus provided effectively for a system which managed, owned, and controlled a number of banks, without infringing on Michigan banking law.

It was this structural characteristic of the COMAC operation which made it distinctive among Michigan banks, for many of the features of the bank holding company were embodied, despite the fact that legally it was not a holding company. However, this was not the only distinctive characteristic, for COMAC engaged in asset and liability management planning much more aggressively than typical midwestern banks. Basically, the asset management plan was designed to maximize profits, regardless of whether the resulting asset portfolio of the bank was conventional, in format. For example, COMAC would select an application of funds, that is, a portfolio application, which represented the best and most profitable use of funds at the margin. This might mean that there would be no accommodation of customer loan requirements during certain periods when it was more profitable to invest bank funds in open market securities, such as municipal bonds, commercial paper, and the like.

This asset management plan runs counter to the traditional position of banks that they should be institutions which serve local market credit needs.

To understand the extent of COMAC’s departure from convention, one must realize that profit maximization has always been a...
difficult problem for bank management. For example, Whitley reports from his survey of bank officers the response to the two following questions: (1) "Is the objective of your bank to serve the public in the manner and spirit of a quasi public corporation?" (2) "Is it the objective of your bank to make as much net profit as long as possible?" The answer to the first question was affirmative among 55 percent of the bank officers queried, while responses to the second question were negative in almost two out of three cases [22]. On the basis of these responses, one is led to believe that a majority of bankers tend to reject profit as the primary goal. Other attempts have been made to assess bank management aims, including a recent study sponsored by the Federal Reserve Bank of St. Louis and conducted by L. C. Anderson and A. E. Burger [23]. The research postulates that if banks are profit maximizers, they will respond differently to changes in market variables than if they are loan accommodationists. The study then attempts to predict how banks will change their preferred holdings of excess reserves, borrowings, and loans under alternative hypotheses. The results suggest that bank management may have begun to exhibit more aggressive behavior both among themselves and toward other financial intermediaries in recent years, but that the "accommodation" principle is still conventional wisdom among many banks.

A second important concept in the COMAC philosophy was the idea that a bank should be aggressive in the retail savings market, the conviction being that if a large proportion of total deposits were in the individual savings category, the bank would have an extremely stable deposit base. This type of deposit was felt to be somewhat insensitive to interest rate differentials, and to the extent that there was a sensitivity, the problem could be minimized by paying maximum allowable interest rates and through use of aggressive marketing techniques. If the above philosophy is correct, two things seem obvious: The strength or need for strong customer relationships is weakened, and high gross income is needed to pay for the high proportion of relatively high cost time money.

The above philosophy has implications for problems of liquidity management. If the requirement to maintain customer relationships is minimized, than the need for liquidity arises from three sources instead of the usual four. Liquidity to meet future loan demands is no longer a large factor, but liquidity is still needed for the uncertainties of future deposit levels, the number of loan defaults, and changes in the value of investments.

Thomas Gies 143

Liquidity is the ability to cope with reductions in asset values and reductions in volume of liabilities. From the asset side, need for liquidity can be accommodated by encouraging short-term investments, maintaining a high quality loan portfolio, and holding large amounts of cash. However, as noted earlier, these policies tend to be costly in terms of profit. From the liability side, a bank can gain liquidity by minimizing deposit volatility or by developing potential sources of funds that can be tapped as occasion demands. Hence, the two basic methods of achieving liquidity are by having assets that are convertible to cash and by having a source of funds that can be developed to replace lost deposits or to meet needs for incremental deposits.

Conventional methods of meeting liquidity needs have concentrated on the first technique. A large proportion of bank assets traditionally have been concentrated in what are called primary and secondary reserves. Twenty years ago nearly 50 percent of all bank assets would have been in the form of cash and government securities. While these very high proportions of cash have been reduced in recent years, it is still true that many banks depend heavily upon a liquid portfolio in order to meet the liquidity requirements. The COMAC philosophy, on the other hand, stressed the concept of a bank's liquidity being met by the simultaneous creation of an asset and a liability. A part of this philosophy was the idea that the holding of municipal securities was not necessary for liquidity purposes. This is an extension of the nonaccommodation posture in portfolio design described earlier. There are two main points in the COMAC approach to liquidity management. First, a bank does not need to maintain a great proportion of short-term liquidity in its asset accounts. Second, a bank's ability to create claims against itself should be used to meet immediate liquidity demands. Placing primary emphasis on this approach frees all but an absolute minimum of cash in governments for pure investment purposes. In addition to these two techniques, COMAC used the more generally accepted tool of liquidity management, asset sales, and participations.

In the earlier discussion it was pointed out that the bank using liability creation for liquidity had to ensure itself of constantly being able to attract new funds, and, because of interest rate ceilings, a bank could not always depend upon being able to do so. Thus, it had to be certain it could obtain loans on the federal funds market. To ensure this source of liquidity, the COMAC banks followed a unique practice in the federal funds market, namely, arranging intermediate term
agreements with nonrelated banks to supply a specified level of federal fund loans. This amounted to a secured line of credit in that market. The COMAC banks paid fees for these lines of credit, but obtained what the banks needed: an assured source of funds to meet liability drains. It became evident through the late 1960s that the technique satisfied a very important need, for, as deposit rates reached Regulation Q ceilings and it became difficult to raise funds in the deposit market, time deposits as a new source of funds for the COMAC banks tended to diminish, and the federal funds market dramatically increased as a source of funds. The volume of such borrowed funds became so important that it truly can be said that some of the COMAC banks were borrowing their total reserve requirements.

In summary, the distinctive characteristics of the COMAC banks were as follows: (1) Loans were a relatively less prominent part of the bank's portfolio; (2) municipal securities, particularly long maturity issues, were an extremely important element; (3) time deposits as opposed to demand deposits were an important source of funds; and (4) the federal funds market was an immensely greater source relatively for COMAC banks than for banks generally. In addition to these asset and liability management characteristics, it should be noted that the general growth in size of the COMAC banks caused their capital ratios to diminish to a level that would be regarded as near the lower margin for Michigan banks.

An intangible factor also needs to be stated, although it is difficult to assess its exact importance in conjunction with the other characteristics of COMAC management. During the period of COMAC activity, the need to cultivate personal relationships with regulatory authorities came to be well known. Most banks adopt a respectful, somewhat restrained, posture toward regulators. Bank examiners, for example, are handled with a good deal of tact and diplomacy. Although a bank may not comply fully with all regulatory requests, whether from the comptroller of the currency, the Federal Reserve, the Federal Deposit Insurance Corporation, or the state banking commissioner, a bank is very careful not to develop hostile relationships. As was previously mentioned, COMAC personnel made few, if any, efforts, to maintain a smooth working relationship with regulators. An Alternative Hypothesis

An alternative explanation has been suggested for the resistance which COMAC provoked. Stating this hypothesis in clear and extreme form is useful in trying to determine where the truth lies.

Thomas Gies 145

(1) COMAC did achieve entry by take-over, with great speed and effectiveness. This process parallels what bank holding companies are able to do in larger markets, although they rarely do so.

(2) COMAC did not replace managers in large numbers, preferring instead to set new policies. Its take-overs were aimed at new behavior, with a minimum of upset to prior arrangements.

(3) Many of the target banks were classic cases of high inefficiency, offering the possibility of dramatic increases in profitability under new policies. It is not clear that any such degree of inefficiency exists widely in larger, metropolitan banks.

(4) COMAC's changes were to shift the portfolio toward state and local bonds and to introduce a small degree of price competition to obtain savings deposits.

(5) The effect on earnings was dramatic.

(6) The effect on other banks was not clear. Given a longer time in which to work, COMAC might have forced more competitive behavior from its rivals, but apparently two or three years is not sufficient.

(7) COMAC was hampered by public regulators, in a series of specific actions.

(8) This intervention was not justified by the economics of the COMAC banks, for generally they were in a good position to survive the 1969 credit crunch.

(9) There are two possible interpretations of the regulators' actions:
   (a) COMAC was so abrasive that regulators sought revenge and a reassertion of their authority along conventional channels.
   (b) Alternative hypothesis: The regulators acted as agents of the conventional banking establishment, to which COMAC was a threat. If so, COMAC could not have escaped by being more diplomatic.

(10) The opportunities which COMAC exploited are not widely available in large-scale banking, although they may exist in many smaller localities.

(11) Take-over can powerfully affect banking behavior.

(12) Unless regulators are at pains to be neutral or even to promote take-over, they will easily slide into the role of stopping the process, or as it may be, the interest of established banks.

The alternative hypothesis is highly negative, and yet it accords with the most acute recent interpretations of regulatory behavior, especially of bank regulation. It is in banking markets that the
protective role of regulation has become perhaps most fully
developed. Such a role may serve valid economic goals of security and
efficient banking, but in this case it did terminate a sound and
efficient experiment in modern banking, one which was using the
traditional market processes to improve performance.

Both hypotheses can be supported by all or most of the facts about
the whole COMAC episode. I prefer the first and milder alternative,
but the second cannot be rejected.

Summary

The origin and development of the difficulties which the COMAC
group experienced during the period 1966–1970 have been described.
At this point a review of the group’s overall philosophy, the success
which it evidently achieved, and its practicability and effectiveness
seems in order.

The cornerstone of the COMAC philosophy was the idea that a
stable base of time deposits could be attracted to the banks. By
making this type of deposit the major source of funds, a bank could
substantially reduce its short-term liquidity problems. Insufficient
data make it impossible to determine whether each bank was success-
ful in this respect, but the banks in the three markets reviewed here
did attract more than their share of total time deposits. It is not
possible to determine what that proportion was of the class of deposits
from which those banks were drawing.

It is hypothesized here that two circumstances occurred with
respect to deposits during the period 1966–1969. First, it seems
reasonable to assume that the banks were initially successful in
attracting the desired deposits and that they were in fact a stable
source of funds. The banks apparently went through the credit diffi-
culties of 1966 with serious problems. Very little use, for example,
was made of the Federal Reserve discount facility. Furthermore, it is
apparent that the banks were able to continue expanding their
sources of funds through the period of tight credit in 1968 and 1969.
Even in 1969, when the credit squeeze was having its major impact on
all U.S. banks, COMAC banks maintained a significant net growth.

The market for large amounts of certificate of deposit money was
a dependable source of funds and provided a stable deposit base as long
as the Regulation Q ceiling provided enough room for aggressive
banks to compete on a price basis. The banks that have been exam-
ined here were mainly price competitors, despite the effort to carry
out the creation of competitive services in other areas. But by late

1969 Regulation Q ceilings were having a very severe dampening
effect upon time deposit growth and, consequently, upon banks such
as the COMAC group, which depended upon price competition to
attract funds. It appears that the cornerstone of the COMAC philoso-
phy may have been correct, but the banks’ inability to adhere to it
absolutely stemmed from the fact that regulation in the form of a
ceiling on deposit rates prevented them from continuing to be vigor-
ous, aggressive competitors in a tight credit period.

Another major aspect of the philosophy was profit maximization
in asset management. Inherent in this approach was the breaking
down of the traditional customer relationship. It seems possible to
hypothesize two results of this asset management plan. First, the
weakened customer relationships made COMAC banks susceptible
to severe deposit drains, and COMAC banks did suffer from heavier
drains in 1969, although this did not seem to have been true in 1966.
Second, it has been suggested that because of the COMAC banks’
investment policy they were able to establish substantial customer
relationships with municipalities. However, when the real test came
in 1969–1970, these relationships proved unstable. Municipal and
state treasuries were hard pressed to meet the rising demands for
funds and grew sensitive to maximizing income at all points. Thus, as
market rates climbed above the Q ceilings, municipal treasurers
could not continue to justify bank deposits. To gain the added income
they had to break the customer relationship which had been estab-
lished with COMAC banks. In all probability this phenomenon
served to increase the rate of disintermediation which the COMAC
banks experienced. The COMAC philosophy specifically indicated
that municipal bonds would not be used as a source of liquidity.

Accordingly, COMAC banks tended to have relatively long muni-
cipal portfolios because the longer end of the term structure afforded
a significantly higher yield. Specifically, if the COMAC banks had
preserved 25 percent of their municipal bond assets in the liquid form,
40 percent for reacting to market changes (that is, pure profit
maximization), and 25 percent invested permanently in long-term
issues, they would have had a distinctly better liquidity position at
the end of 1969 and in early 1970. It is not possible to say that this
policy would have enabled the COMAC banks and the COMAC group
to survive the first half of 1970, but it would definitely have relieved
some of the pressure on borrowings.

There is nothing to suggest that borrowing short-term liquidity
(liability management) was incorrect. The facts show no problems
with the approach until about March 1970. The technique was obviously the immediate cause of the COMAC debacle, but this was mainly because other parts of the philosophy had proved unworkable in their implementation. There are certain macroeconomic problems if the approach were to be used universally, of course, but this is not to say that it cannot be used for an individual bank or banking group.

References


We have tried to judge how far potential competition and new entry can supplant the orthodox regulatory process. The group has examined both the general attributes of entry as a constraint and the specific prospects for entry into three quite different "utility" sectors: electricity, communications, and commercial banking.

The gist of our appraisals is negative; entry offers little prospect for constraining utility behavior as a regulatory substitute. It can yield beneficial effects in certain secondary markets under properly arranged conditions, but it should not be regarded as a constraint which can simply be invoked while regulation is withdrawn. Instead, it must be carefully monitored, even made, by public policy choice if in fact it is to influence significant utility markets toward efficiency and equity.

This view derives from a variety of specific points, which can be condensed into five propositions. These are provisional, of course, and can be regarded as an agenda for further research.

First, entry is a complex matter. There are difficulties in defining it, measuring it, and in predicting its consequences. There are several elements to entry — including its extent and speed — which require more rigorous treatment. The two main types of entry, by new capacity and by merger, need more analysis and comparison. Neither gross
Summary

Michael Klass and William Shepherd

nor net entry has been measured precisely in the literature, nor have they been distinguished from the process of changes by existing firms. Just how entry affects dominant firms remains unclear: via expectations, the drop in market share, or other causes? Empirical tests of entry's effects have been largely begun, and thus far the results are inconclusive. Therefore, entry could have importance, but its actual nature is uncertain and its role is unproven.

Second, each utility sector has its own specific entry possibilities, which differ in kind and scope. In electrical power, entry would have vertical attributes and arise from mutual penetration of service areas; in telephones, technology is critical to entry; in banking, entry requires diluting the accustomed bank-client relationships. In some sectors, entry has little or no possible scope; in others (truck ing, for example), its potential scope is large.

Third, most potential entry is small and peripheral to the main utility markets. Despite predictable exaggerations by utility spokesmen, new entry poses no severe threat to core activities in communications and power. Even a "fair" entry policy would scarcely affect the structure and profitability of some of the main regulated utilities.

Fourth, genuine open entry is, fundamentally, incompatible with conventional regulation. It involves changes which conflict with the franchise basis of regulation. That basis can perhaps be changed by shifting to franchise renewal processes or by revising the franchise in other ways, and some of these options offer interesting alternative forms of entry. Still, it is often self-contradictory and meaningless to appeal to regulators to foster entry (beyond trivial cases), for it is simply not within their framework of powers, skills, and motivations. To obtain substantial open entry requires legislative actions, which the franchised firm will endeavor to make prolonged and difficult.

Fifth, even where open entry is made possible, it requires specific regulatory actions. Entry must usually be made by regulators, not just permitted. This is because regulation makes the regulators also responsible for service and identifies them with the established suppliers' interests. Furthermore, the regulatory actions to foster entry must be of high quality and sophistication, which exceeds the normal capacities of regulators in the past. One should be skeptical that regulators will handle entry ably, even apart from the other impediments to bringing it about.

In short, the conventional objectives of an open-entry policy, as part of deregulation, appear to be largely unobtainable except in trivial or odd cases. More, not less, independence and quality in regulatory policy choices is required than has been customary in the past.

If such conventional treatments will give meager yields, one turns to other, less obvious alternatives. These include such strategies as franchise renewal processes, partial public ownership and direct control, take-overs as a source of managerial constraints, vertical restructur- ing, public enterprise competitions, and others. Each deserves extensive treatment. For now, we have explored entry as a substitute for regulation and must report that it appears to be, as is true of so many other promising ideas, more hopeful as a theoretical possibility than as a practical one.
Index

"Above 800 Megacycles" case, 98
Advertising intensity, 46 - 48
Algonquin-Northeast pipeline, 70
Alhabed, David A., 125
Aluminium industry, 56
American Gas Association, Inc., 72
Arnott Interconnet Companies, 102
Armour & Co., 51, 56
Asset management plan, 111
Anderson, L. C., 142
AT&T, 94 – 103, 107 – 14; see also Bell Telephone system
Aviation for managerial functions, 59
Automobile industry, 51
Averaging principle, 69
Averch, Harvey, 9 – 13, 17, 19 – 32, 25 – 30, 31, 32, 106
Backman, Julian, 50
Ball, Elizabeth E., 23, 26, 27, 32
Banks, Joe E., 2, 26, 39, 40, 41, 43, 44, 46
Bank market organization, static versus dynamic, 124 – 27
Bank of the Commonwealth, 139
Bank of Montana, 135 – 36
Bank management, criticism of, 128 – 30
Banking markets, 118 – 19; empirical studies, 125 – 26; regional and local, 21 – 24
Banking, price and output determination, 119 – 25
Barriers, incentives to enter and the forms of regulation, 17 – 21; role of in large-scale industry, 42 – 50; profitability, 43 – 48; man-made, 53
Baumol, William J., 9 – 11, 25, 32
Bell Telephone system, 14, 58, 94 – 103, 107 – 12
Berry, Charles H., 49, 50
Bewkes, D. A., 104
Burger, A. B., 142
Carterfone device, 99; decision, 99
Caves, John A., Jr., 71
Chamberlin, E. H., 38
Chandler, Lester, 117
Chandeyre, Robert, 125 – 26
Clark, John Bata, 37 – 38
Coca-Cola Co., 51
Coffam, Robert, 26, 27, 32
COMAC, 130 – 35; and the regulatory agencies, 139 – 44
Campbell, James S., 55
Capron, William M., 58
Civil Aeronautics Board, 6
Consol, William S., 46
Common carrier status: reality, 81 – 83; concept, 83 – 85
Computer industry, 51
Comsat, 114
Cooke, Louis, Jr., 70
Cross-subsidization, 6 – 7, 89, 98
Crude oil industry, 63 – 69; transportation, 63 – 64
Datran, Inc., 100, 102
Dellott, John D., 94, 114
Demsetz, Harold, 14, 40
Department of Justice, 83, 1941 consent decree, 89
Dewey, Donald J., 28, 30, 49
Dow Chemical Co., 133
Douglas, 114
Eastman Kodak Co., 53
Economies of scale in the telecommunications industry, 104
Ermann, Peter C., 133
Electronics, 96
Electrical goods industry, 51
Electricity industry, 6, 63 – 69, 73 – 91; entry options, 73 – 76; load factor, 79; nondiscriminatory access of distributors, 91; power pool, 79 – 81; regulation problems and solutions, 85 – 89; restructuring, 74 – 76; serial entry, 76 – 78
Entrance profiles, 58
Entry: absence of entry into high profit industries, 55 – 56; barriers, 56; benefits, 56; control, 12; device limiting entry, 54; and the rate of return, 22 – 26; legal prohibitions, 17 – 18; losses and effects, 51 – 52; into management,