

# Taxation by regulation

Richard A. Posner

Professor of Law  
The University of Chicago

*Students of the regulated industries often assume that regulation is designed either to approximate the results of competition or to protect the regulated firms from competition. But neither view explains adequately a number of important phenomena of regulation and regulated industries. Foremost among them is the prevalence of “internal subsidies,” whereby unremunerative services are provided, sometimes indefinitely, out of the profits from other services. To understand this and other phenomena, we must assign another important purpose to regulation: we can call it “taxation by regulation.” The purpose of this paper is to explore this dimension of the regulatory process, to demonstrate that it explains some otherwise perplexing features of the process and the industries subject to it, and to compare it with other methods of public finance.*

■ Two views of the purpose in fact of public utility and common carrier regulation<sup>1</sup> vie in current scholarly debate. One, the more familiar, holds that regulation is a device for protecting the public against the adverse effects of monopoly<sup>2</sup>; the other holds that regulation is procured by politically effective groups, assumed to be composed of the members of the regulated industry itself, for their own protection.<sup>3</sup> In my opinion neither view, at least as thus far formulated, explains an important phenomenon of regulated industries: the deliberate and continued provision of many services at lower rates and in larger quantities than would be offered in an unregulated competitive market or, *a fortiori*, an unregulated monopolistic one.

---

Richard A. Posner, who received the A.B. degree from Yale University and the LL.B. from Harvard, is Professor of Law at the University of Chicago Law School. Among his major interests is the application of economic theory to law, and he also is the editor of a new journal to be published under the auspices of the University of Chicago. It is entitled *The Journal of Legal Studies*, and its first appearance will be in January 1972. The author would like to thank Edmund W. Kitch and George J. Stigler for their helpful comments on an earlier draft.

<sup>1</sup> By the awkward term “purpose in fact” I mean to distinguish sharply between the reasons ascribed to regulatory laws and decisions by the legislators and regulators themselves and the reasons, whether or not anywhere avowed, that provide a consistent explanation of the actual course and consequences of regulation. It is with the latter that I am concerned here.

Although the emphasis throughout this paper is on the full panoply of rate and entry controls characteristic of public utility and common carrier statutes, where the analysis can be extended to markets having somewhat different regulatory arrangements I have not hesitated to do so.

<sup>2</sup> For example, see Bonbright [3], p. 23.

<sup>3</sup> A rigorous statement of this theory appears in the article by G. J. Stigler in this issue of *The Bell Journal* [59]. The “capture” of regulation by the regulatees is, of course, an old theme in the literature of regulation. Professor Stigler’s theory allows for capture by effective political groups other than the regulated firms themselves, and there is accordingly no necessary inconsistency between it and the analysis in this paper.

This phenomenon can be explained, I believe, only if we modify existing views by admitting that one of the functions of regulation is to perform distributive and allocative chores usually associated with the taxing or financial branch of government. And we shall see that an analysis of taxation by regulation explains other perplexing phenomena. But it would be error to think that the analysis compels rejection, as distinct from modification, of the existing views of regulation. I hope to show that any theory that conceives the function of regulation to be to approximate the results of competition, or to enrich the regulated firms, or to do sometimes the one and sometimes the other, is incomplete. But it does not follow that a broadened public-interest approach (one that accommodated certain subsidy elements) or a broadened effective-political-group approach (one that viewed certain customer classes as effective political groups) might not be tenable.

■ The best known example of a regulated service that an unregulated competitive market would not provide on the same scale is railroad passenger service,<sup>4</sup> but it is far from being an isolated example.<sup>5</sup> Like railroad passenger service, domestic telegraph service would be declining even more rapidly than it is, were it not for the Federal Communications Commission's stubborn rearguard action against further rate increases and service degradation by Western Union.<sup>6</sup> Local airline service provides another example,<sup>7</sup> complicated though by the presence of direct as well as internal subsidies.

The phenomenon is by no means limited to declining industries. The American Telephone and Telegraph Company may soon be required to provide electronic interconnection free of charge (or nearly so) to the National Educational Television network.<sup>8</sup> Broadcasters are required to provide at least some nonremunerative news and public-affairs programming.<sup>9</sup> Liability insurance for high-risk automobile drivers in many states is written at a loss.<sup>10</sup> Producers of natural gas are constrained to sell at a price that does not include scarcity rents, thereby benefiting present consumers of natural gas at the expense of future consumers, who may encounter shortages.<sup>11</sup> Uniform rates, based on averaging together the costs of services whose cost characteristics are in fact very different, are a conspicuous feature of regulated rate structures. Examples are statewide telephone rates, the uniform long-distance rate for telephone calls of the same distance and duration regardless of route, the pegging of airline and rail passenger rates to distance with little or no consideration for cost differences over different routes, and the flat rates characteristic of

---

<sup>4</sup> See Conant [10], p. 132, Hilton [23], p. 136, and Nelson [40], pp. 286–301.

<sup>5</sup> In at least one respect, we shall see, it is atypical of the general class. See p. 43 and fn 69 *infra*.

<sup>6</sup> See President's Task Force on Communications Policy [50], and U. S. Federal Communications Commission [75].

<sup>7</sup> See Eads [14], pp. 1, 13–14.

<sup>8</sup> See The Communications Act of 1934 [78], sec. 396(h), Interconnection Service [73].

<sup>9</sup> See Public Service Responsibility of Broadcast Licensees [74].

<sup>10</sup> See Keeton and O'Connell [30], pp. 93–94; cf [67], Price Variability in the Automobile Insurance Market.

<sup>11</sup> See MacAvoy [35], p. 271.

## 1. Internal subsidies

urban transit systems.<sup>12</sup> Electrical companies grant discounts, unrelated to any cost savings, to hospitals and other worthy groups,<sup>13</sup> while water companies are often required to furnish water for household use at rates below marginal cost—and to fire departments and schools for nothing.<sup>14</sup>

These practices are very old. In 1827, the State of Illinois required by statute that every keeper of a ferry, toll bridge, or turnpike road “give passage to all public messengers and expresses; to all grand and petit jurors, when going to and returning from court, without any fee or reward whatever” [26].

In all of these examples, as in many others one could cite (including some to be discussed later), a service is provided that does not pay its way in the market. Someone must pay its way, however. Normally it is other customers of the firm rendering the service, who pay a price higher than the cost of serving them. Sometimes it is future generations of consumers (as in the gas example) or the stockholders or creditors of the firm (as, perhaps, in the railroad and telegraph examples). Examples of internal or cross-subsidization, as we may call the practice, lie everywhere at hand in the regulated industries.<sup>15</sup> They are also commonly found among public enterprises here and abroad,<sup>16</sup> the structure of postal rates providing a conspicuous example.<sup>17</sup> They seem much less common in unregulated private markets. The contrast is instructive: it suggests that the provision of internal subsidies is associated with distinctively governmental purposes and functions. We shall return to this point.

Before proceeding, I should caution the reader that identification of our phenomenon—internal subsidization—is often difficult.<sup>18</sup> In practice an internal subsidy is not always easy to distinguish from certain familiar profit-maximizing practices. A monopolist may be able to maximize profit by setting different markups on different sales, depending on the elasticities of demand of particular customers or groups of customers. Like the internal subsidy, this practice involves a departure from cost-based pricing. But it does not involve any unremunerative services. Were the lower-priced services unprofitable, the monopolist could increase his profits still further by terminating them. To take another example, a firm that has a monopoly in some markets can, under certain restricted conditions, increase its profits by selling below cost in other markets where it still faces competition. Unlike internal subsidization, however, such “predatory price discrimination” is strictly a temporary tactic to

---

<sup>12</sup> See Caves [5], p. 369; Davidson [12], p. 217; Johnson [27]; Garfield and Lovejoy [18], pp. 200–203, 243; Meyer, Kain and Wohl [37], pp. 354, 357; Meyer Peck, Stenason and Zwick [38], pp. 166–67; Watkins [82], p. 623.

<sup>13</sup> See Bonbright [3] p. 111.

<sup>14</sup> Hirshleifer, DeHaven & Milliman [24], pp. 109–11; Garfield and Lovejoy [18], p. 225; cf. Keig, Fristoe and Goddard [31].

<sup>15</sup> For additional examples see Bonbright [3], pp. 111–12; Friedlaender [15], pp. 66–68; Meyer, Peck, Stenason and Zwick [38], pp. 194–95; Nelson [40], p. 331; Hearings on Transportation Acts Amendments [66].

<sup>16</sup> See, for example, Coase [8], pp. 1, 13–14; Crew [11], p. 258; Peltzman [44]; Sargent [51], p. 248; Sharp [52], p. 53; Shepherd [53], p. 132. Sargent and Shepherd analyze internal subsidization in the context of public enterprises in terms somewhat parallel to mine in this paper.

<sup>17</sup> See Baratz [2], p. 305; Coase [9], p. 25; Kennedy [29], pp. 93–94.

<sup>18</sup> As emphasized by Nove [42], p. 847.

drive out competitors; prolonging it indefinitely would make no business sense.

Peak-load pricing also involves different prices for the same service, and is also different. One observes resort hotels charging, for identical accommodations, higher prices in peak seasons than in off-seasons. But the reason for the off-season discount is that it attracts customers who would otherwise not buy the service. So long as the price they pay covers the marginal or additional cost of serving them, the discount is profitable. It is therefore not a case of internal subsidization as I use the term.

Temporary below-cost selling might of course stem from errors of various kinds. And even persistent departures from a proper matching of price and cost could reflect simply the difficulty and expense of determining which costs were incurred in making which sales; some averaging of costs over different services is certainly explicable on this ground. Other cases of apparent below-cost selling evaporate on examination. Professor Bonbright gives as an instance of internal subsidization (he calls it “social principles of rate making,” but it is the same concept) the granting of special rail fares to clergymen<sup>19</sup>; but one suspects that such a discount is simply a public-relations gesture. In view of the low income of clergymen, which may make their demand for some services relatively elastic, the discount might even be an example of price discrimination.

Finally, some services may appear unremunerative only because an inappropriate conception of cost is employed. As the resort example shows, a price below average cost, but equal to or above marginal cost, may still be remunerative. On the other hand, a price above even long-term marginal cost may be unremunerative.

The last point requires some explanation. In the case of an industry in which average cost decreases with output, a firm that charged a uniform price equal to its marginal cost would not recover its total costs. It could recover them by setting a uniform price equal to average cost. This would force customers willing to pay a price equal to or slightly above marginal cost but not the higher price equal to average cost to turn to more costly substitutes. Neither result is optimal, and the proper solution to the dilemma is a matter of fair debate. One attractive possibility is to charge a price equal to marginal cost for marginal purchases and a sufficiently higher price for inframarginal purchases to cover total costs without losing those sales. Although the proper design of the rate structure is not easy, this approach seems preferable to either the uniform marginal-cost price, which necessitates a government subsidy to make up the deficit in covering total costs, or the uniform average-cost price, which excludes the marginal sale.<sup>20</sup>

It is implicit in this solution, however, that some customers may end up paying a higher price than under average-cost pricing. A simple arithmetical example will illustrate. Suppose the fixed costs of providing a service are \$1000 and marginal cost is \$1, and suppose further that when price is made equal to average cost 500 units are demanded and supplied at a price of \$3 per unit. Therefore, Buyer

---

<sup>19</sup> See Bonbright [3], p. 61.

<sup>20</sup> See Coase [7]; Henderson [22], p. 223; Hirshleifer, DeHaven and Milliman [24], pp. 91–92.

X, who purchases 20 units, pays a total price of \$60. Now suppose the industry adopts a two-part pricing system (one version of the preferred solution to the dilemma of proper pricing under conditions of decreasing average cost) under which anyone wishing to purchase service must pay (a) a share of the fixed costs (ideally, the shares should vary inversely with the elasticity of the buyers' demands) plus (b) the marginal cost of each unit he buys. And suppose that X's share of the fixed costs is set at \$100. He must now pay \$120 for the same 20 units.

To the extent that regulatory agencies are unwilling to visit such consequences upon particular customers of the firms they regulate, output is not carried to the efficient point. The subsidy is indirect but inescapable: the additional price that the rejected marginal customers must pay for substitute products is a cost imposed on them in order to enable inframarginal customers to buy at a cheaper price than if an efficient pricing system were employed. Profit-maximizing price discrimination would also result in the inframarginal customer paying a higher price. The pricing actually employed in decreasing-cost industries appears often to be neither efficient nor profitable.

"Value-of-service" pricing in the railroad industry provides good examples of this point. Before the development of truck transportation, the practice of proportioning rail rates to the price of the commodity transported may have been a roughly adequate method of concentrating the fixed costs of railroad service on those customers whose demands for rail transportation were least elastic. But once there was a good substitute service, the method ceased to have a rational basis, at least under the usual views of regulation, since the price of the commodity shipped bears no necessary relation to the adequacy of trucking as a substitute for transportation by rail. Wristwatches are more expensive per unit of weight than grain, but it does not follow that the demand of the wristwatch shipper for rail service is less elastic; it is probably more elastic, because trucking is a more feasible alternative for the shipper of wristwatches than for the shipper of grain.

Value-of-service pricing may have persisted because it is a convenient method of subsidizing some shippers regardless of the elasticity of demands for rail transportation. The favorable rates at which agricultural commodities continue to be transported seem a case in point.<sup>21</sup> Considering the broad range of subsidies that farmers have managed to obtain for themselves, it is perhaps not surprising that they have obtained internal subsidies as well. And they are not the only group, or railroading the only industry, where we find value-of-service pricing used as a method of internal subsidization.<sup>22</sup> Notice that keeping rates low to customers so favored is not only inefficient but can destroy a decreasing-cost industry.<sup>23</sup> These consequences only underscore the limitations of existing theories of regulation.

A final problem of characterization, one that also arises from the phenomenon of decreasing average costs in some regulated industries, requires brief mention. A pricing system for a decreasing-cost in-

---

<sup>21</sup> See Friedlaender [15], pp. 18–20; The Hoch-Smith Resolution [80], p. 801; cf Meyer, Peck, Stenason and Zwick [38] pp. 175, 187–88.

<sup>22</sup> See, for example, Johnson [27], pp. 42, 67; Wilson [83], p. 337.

<sup>23</sup> Stigler [60], pp. 213–14.

dustry, to be efficient, must satisfy three criteria: (1) it must enable the total costs of the enterprise to be recovered; (2) it must be so designed that no customer willing to pay at least the marginal cost of serving him is turned away; (3) there should be no sales below marginal cost. Now there are many different formulae for allocating overhead costs consistently with these criteria and the choice of one inescapably affects the distribution of wealth as between various customers. For example, suppose the fixed costs of an enterprise are \$100; there are 10 customers who can be made to contribute to those costs; and each of them would pay \$20 rather than do without service. The regulatory agency can insist on a pricing scheme under which each pays \$10, or on one under which some pay \$20 and some less (perhaps zero). Whichever choice is made makes some customers worse off and some better off than they would be under an alternative arrangement.

Because this type of “subsidization” does not involve the maintenance of any uneconomic service, and thereby avoids most of the problems with which this paper is concerned, I exclude it (perhaps arbitrarily) from the category of internal subsidies. That regulation unavoidably involves issues of wealth distribution as between customer classes does, however, emphasize the relevance of notions of taxation and public finance to the theory of public regulation—but on this, more shortly.

■ The existence of the internal subsidy is an embarrassment to proponents of the first view of regulation (at least as usually formulated) mentioned at the beginning of this paper—the view that regulation is imposed in order to bring about results approximating those of competition. As we have seen, the internal subsidy brings about results unthinkable in a competitive market, which is perhaps why a distinguished proponent of the first view, Professor Bonbright, who believes that “public utility services are designed to be sold at cost, or at cost plus a fair profit”<sup>24</sup>, considers the internal subsidy aberrational.<sup>25</sup>

Nor is it explicable if regulation is conceived purely as a service demanded by and supplied to the regulated firms themselves, although here the demonstration is more complicated. While an unregulated firm, whether monopolist or competitive, would not profit from engaging in internal subsidization, regulation alters normal business incentives. Under certain conditions, a regulated firm can increase its profits by expanding the size of its plant even if some of the output of the enlarged plant must be sold at a loss.<sup>26</sup> In addition, a regulated firm might provide some unremunerative services as an unintended byproduct of a policy of deliberately not keeping detailed cost information. The purpose of such a policy would be to prevent the regulatory agency, which is normally dependent on the regulated firm for information about the firm and its markets, from acquiring sufficiently detailed knowledge of the firm’s operations to enable more stringent regulation. Such a strategy is less risky than it may seem. A regulated monopolist can get by with less information about

## **2. Internal subsidization and the received views of regulation**

---

<sup>24</sup> See Bonbright [3], p. 23.

<sup>25</sup> *Id.*, Ch. 8.

<sup>26</sup> Averch and Johnson [1], p. 1052.

the costs of particular sales than a nonregulated firm. It may be able to earn the overall profit permitted by the agency, even though many of its sales are unprofitable, by charging supracompetitive prices in markets where its monopoly is secure.

Furthermore, a firm that engages in internal subsidization can argue forcefully to the regulatory agency that the agency should not permit, or at least should strictly limit, the entry of competitors into those markets where the firm makes large profits, because those profits—which new entrants would erode—are necessary in order to cover the losses in the subsidized markets. Finally, regulated monopolists may believe that differential rates, even where cost justified, strike the public as discriminatory and may in consequence invite sterner regulation of the firm's profits.

These considerations do not explain, however, those instances, such as railroad passenger service and natural gas production, where the regulated industry does not appear to be benefiting, directly or indirectly, from internal subsidization. Nor do they explain very convincingly why the practice is ever instituted in the first place. If regulated firms dominate an agency, as the view under examination posits, they can use their position to increase profits directly by getting the agency to fix a high permitted level of profits and forbid the entry of new competitors. A program of internal subsidies seems a needlessly roundabout method for achieving these ends, and also a costly one: it entails smaller profits for the regulated firms since, as mentioned earlier, a profit-maximizing firm is always better off terminating a losing service. If firms do not yet dominate a regulatory agency but seek to do so, a program of internal subsidies is not likely to help them to reach their goal. A regulatory agency that is not a creature of the firms being regulated will not permit plant used for rendering unremunerative services to be figured in the rate base. It will not reason from the fact that the regulated firms are losing money in some markets to the conclusion that it should forbid the entry of new firms into others; rather, it will invite the firms to terminate the losing services. It will insist that the regulated firm adopt accounting procedures that reveal the true costs of serving various classes of customers. These regulatory checks may be far from entirely effective in practice, but it does seem unlikely that, if regulatory agencies had no independent reasons for encouraging internal subsidization, regulated firms could nonetheless engage in the practice on the scale they do.

In sum, existing views of regulation do not explain well the important phenomenon of internal subsidization. A new approach is needed and, in the next part, proposed.

### **3. Internal subsidization as a branch of public finance**

■ Regulation and public finance are ordinarily considered unrelated activities. Occasionally the language of one laps over to the other, and we speak of a monopolist “taxing” his customers, redistributing wealth from them to him, by charging a price above cost. The internal subsidy, it seems to me, is an aspect of public finance in what is at once a more exact and a more natural sense. Taxation in common parlance refers to the use of the powers of the state to extract money from its subjects in order (1) to defray the cost of services that the politically dominant elements of the state wish to provide and

that the market would not provide in the desired quantity and at the desired price, or (2) to transfer money from one group to another, or (3), often, to do both. By this test regulation is in part a system of taxation or public finance. The basic mechanism is the internal subsidy. A firm provides a service below its real cost, and the deficit is made up by (usually) other customers of the firm who pay higher prices than they would otherwise. Were it not for the power of the state, acting through the regulatory agency, to control entry, the system would not be viable. A firm would not institute a losing service. If by mistake it did, it would terminate the service rather than subsidize the losses from profits in other markets. If it foolishly persisted, firms not burdened with the costs of losing services would enter the high-profit markets and their competition would drive down the price; deprived of the necessary supernormal profits, the firm would finally be compelled to terminate the unprofitable service.

Internal subsidization may thus be viewed as an exertion of state power whose purpose, like that of other taxes, is to compel members of the public to support a service that the market would provide at a reduced level, or not at all. It is in fact a form of excise tax, with the burden falling on purchasers of certain goods or services, and the proceeds earmarked for specific uses. As a form of excise tax, it invites comparison with other methods and objects of taxation.

■ It may be helpful at this point to illustrate the thesis that regulation is a method of public taxation and expenditure by some cases. I confine myself to two: the wall of protection thrown up around the international telegraph carriers by the Federal Communications Commission, and the methods of regulation used in the cable television industry. These case histories are no more than suggestive, but I believe a broader study of regulated markets would confirm the picture they present.

□ **International telegraph service.** Since this is one of the less well known regulated industries, it may be well to begin with some background.<sup>27</sup> U. S. companies providing international telegraph service (which includes telegrams, teletype, and transmission of computer data) are regulated by the FCC and known as “record carriers,” because they transmit communications in the form of a written or otherwise stored record, as distinct from voice-only telephone transmission. There are three principal record carriers: ITT World Com, RCA Global Communications, and Western Union International. Until the 1950’s they owned and operated submarine telegraph cables and high-frequency radio transmitting and receiving equipment. AT&T owned high-frequency radio equipment too, with which it supplied overseas telephone service (underwater cables lacked sufficient capacity for such service). AT&T did not offer record services. It seems to have been understood that the FCC would not permit the telephone company to enter the record market in competition with the record carriers.

A technological revolution in the provision of international telecommunication service occurred in the 1950’s with the perfection of

---

<sup>27</sup> A more detailed description of the industry is given by the President’s Task Force on Communications Policy in [49] and [50].

#### 4. Two case studies

the voice-grade submarine cable, developed and installed by AT&T. These cables had sufficient capacity not only to enable telephone communication but also to permit the derivation of telegraph circuits at much lower cost than was possible to the record carriers using their existing, and now largely obsolete, equipment. Overseas telephone calls were cheaper and of much better quality than before, and since a telephone call is a substitute for a telegram or other record service, the new development placed a good deal of pressure on the record carriers. In addition, AT&T began to offer a new service, called "AVD" (alternative voice-data), under which the customer leased a circuit in the AT&T cable which he could use for voice or record service or both. Leased lines represented the most lucrative portion of the record carriers' services, and AT&T's new offering posed a serious threat to those carriers' continued profitability and even to their survival.<sup>28</sup>

They complained to the FCC. The Commission forbade AT&T to offer further AVD service (although AT&T was permitted to continue serving a few markets where the service had already been instituted) and granted the record carriers rights of co-ownership in AT&T's cable.<sup>29</sup>

The record carriers thus weathered this storm, but another was brewing. In 1962 the Communications Satellite Corporation was created to exploit the newly developed communications satellite. The enabling legislation<sup>30</sup> was ambiguous as to whether Comsat was free to deal directly with communications customers in addition to leasing circuits to the record carriers and AT&T. The issue came to a head when Comsat and the Department of Defense negotiated for the lease of thirty circuits in Comsat's Pacific satellite to the Department directly. Comsat offered a rate that was 32-to-40 percent of the rates offered by the record carriers.<sup>31</sup> The disparity was puzzling, since on leased circuits the record carriers do little more than arrange for obtaining a circuit either in one of AT&T's submarine cables, or, in the present case, in one of Comsat's satellites. After the record carriers agreed on a general rate reduction, although not to the level offered by Comsat, the Commission held that as a matter of policy (not law) it would forbid Comsat to deal directly with the communications user.<sup>32</sup>

Bare as this summary is, it indicates the salient fact of regulation in the international communications industry: the insulation of the record carriers from direct competition. AT&T could have competed directly with the record carriers before the invention of the modern submarine cable and especially afterward, as could Comsat, but the regulatory agency prevented such competition. The received views of regulation do not explain well this persistent, indeed dogged, policy of regulatory protection.

If regulation is imposed in order to prevent monopolistic distortions, one must find reasonable grounds for thinking that the entry of AT&T or Comsat into the record business would have reduced

---

<sup>28</sup> See FCC proceedings re the American Telephone and Telegraph Company [68], p. 1159 and [70], p. 433, n 9.

<sup>29</sup> Re The American Telephone & Telegraph Co. [68].

<sup>30</sup> Communications Satellite Act of 1962 [79].

<sup>31</sup> President's Task Force on Communications Policy [49], p. 19.

<sup>32</sup> *Authorized Entities and Authorized Users* [72].

rather than increased competition in that industry. Only if one attributes simplistic, although admittedly common, notions of competition to the FCC is it possible to construct a competitively based rationale for the denial of entry. The argument would be that AT&T controls the domestic leg of most international record traffic, and if permitted to offer international service as well would use its domestic position to monopolize the international service by routing all traffic to itself.

The fallacy in this reasoning lies in the assumption that a monopolist of one stage in the distribution of a good can increase his monopoly return by annexing successive stages as well. The assumption is in general false.<sup>33</sup> The successive stages of distribution are exactly analogous to the sale of complementary products, for example the head and shaft of a hammer. A monopolist of hammer heads could not increase his profits by getting control of the production and sale of the shafts as well. Any monopoly profits to be earned from controlling the manufacture of hammers could be captured by control of one essential component of hammers such as the heads. A monopolist's interests are ordinarily best served by minimizing the cost of complementary products, which will usually require the encouragement of competition in their provision. Thus, if AT&T indeed enjoys an effective monopoly of the domestic leg of international telegraph service, it should be able to extract all possible monopoly profits by means of appropriate charges for that leg. It would have an incentive to enter the international end of the business only if its costs in providing that service would be less than the rates charged by the existing carriers, in which event permitting it to enter would further the ends of competition.

Even if fear of AT&T's monopolizing the international business were plausible, it would not explain why the Commission also forbade Comsat to enter, or why the record carriers' rates at the time of Comsat's attempted entry were more than twice their costs.

The interest of the record carriers in repelling new entry is of course evident, but it seems unlikely that the protective measures taken by the Commission were merely in response to their demand. Although the determinants of political power are unfortunately not clearly understood, one is skeptical that the record carriers by themselves could have induced the Commission to subordinate to their protection the interests of AT&T and the Department of Defense, two of the nation's largest institutions, and of Comsat. The record carriers appear to lack the most important constituents of political influence. Their total revenues in 1968 were \$173 million, in contrast to the Bell System's total earnings of \$2.4 billion.<sup>34</sup> The record carriers have few employees<sup>35</sup> to go to bat on their behalf,

---

<sup>33</sup> See Bowman [4], p. 19. This statement and the analysis that follows in the text do not hold in a few special cases. Forward integration might facilitate price discrimination, enable the transfer of profits to an unregulated affiliate, forestall new entry, or simply complicate regulation of the primary market. None of these seem likely factors in the case under discussion.

<sup>34</sup> See Federal Communications Commission Annual Report [71].

<sup>35</sup> I am informed by the FCC that as of October 31, 1969, the international record carrier industry (the three major record carriers plus four smaller ones) had a total of 5623 employees in the United States and another 2432 abroad, of whom an undisclosed number are foreigners.

and the carriers' operations are confined to a handful of major cities—principally New York, San Francisco, and Washington<sup>36</sup>—in whose economies they play an insignificant role. One can challenge the comparison with Bell. Since two of the record carriers are parts of very large firms,<sup>37</sup> perhaps the proper comparison is between the parents' revenues (employees, etc.) and Bell's, or between the record carriers' revenues and those revenues of the Bell System that are earned in the international market alone. What this point ignores, however, is that virtually all of the Bell System's operations (even those regulated by state agencies) are directly or indirectly subject to control or influence by the FCC. The entire system, therefore, has a stake in decisions affecting a part. A decision limiting competition by AT&T in the international market could have precedential force in cases involving its right to compete in other markets. The common carrier interests of RCA and ITT, in contrast, are limited to the international telegraph industry.<sup>38</sup> The impact of regulatory decisions in that industry on the parent companies of the record carriers is consequently much smaller than the impact on the Bell System, so one would not expect them to devote comparable efforts to prevailing before the Commission.

A public finance view of regulation provides a clue to the curious success of the international telegraph industry in insulating itself from competition. The record carriers have long provided telegram service at a loss which they recoup by charging supracompetitive prices for other services, principally leased lines, where their costs, as mentioned, are small.<sup>39</sup> As a result, there is a class of customers who receive a service for which they would have to pay much higher prices were it provided in a free market; possibly the service would not be offered at all. These customers would be injured by any policies, such as free entry, which jeopardized the continued provision of international telegram service at the present attractive rates. They constitute allies of the record carriers in seeking the protection of the Commission against new entry and help to explain why such protection has been obtained in the face of strong opposition.

One would like to have a clearer idea of who the class of benefited customers is and how important cheap international telegraph service is to them. We can speculate that they are mostly small firms and individuals—travel agents, some importers, many tourists and their families—rather than large firms: a large user of international telecommunications service would lease circuits or at least subscribe to teletype service. But we know from other industries, notably the retail-drug industry, that small firms (when many in number) are frequently an effective political group in obtaining protective legislation for themselves. And even if the favored telegraph customers do not constitute an effective political group, perhaps they are viewed for one reason or another as particularly deserving, and on that ground favored. A careful empirical study might refute the suggestion that the interests of this customer class are a significant factor in the

---

<sup>36</sup> In re American Telephone & Telegraph Company [68], p. 1158.

<sup>37</sup> Western Union International is not affiliated with the domestic Western Union Company.

<sup>38</sup> ITT owns some South American telephone companies, but they are, of course, not regulated by the FCC.

<sup>39</sup> See *Authorized Entities and Authorized Users* [72], pp. 432–35.

actions of the Commission, but the important point is that such a study seems clearly indicated.

□ **Cable television.** The cable television industry,<sup>40</sup> in both its economic and its regulatory characteristics, resembles the public utility industries of the late nineteenth century. The industry is in its promotional phase. The systems already installed may be only a small fraction of what we can expect when it is mature. As with the earlier public utility industries, it appears that the returns to scale are large, and it may be inefficient for more than one cable television system to serve any given local area. There is a growing momentum of regulation. All municipalities now require providers of cable television service to obtain a municipal franchise; the franchises are becoming more elaborate; and some states are beginning to regulate the rates charged by cable firms.

On the basis of the received views of regulation, one might hypothesize that regulation had been imposed because the service was thought to be a natural monopoly, or had been demanded by the cable companies themselves in order to create or entrench a monopoly position. But neither hypothesis, separately or in combination, explains the specific types of regulation that have been imposed. On the one hand, the rates charged by the cable companies to their subscribers are rarely regulated (although, as mentioned, there is some movement in that direction); the thrust of regulation thus far has not been to eliminate monopoly pricing. On the other hand, municipal authorities have required cable franchisees to pay substantial fees, in money and kind, to the municipal government. These commonly take the form of the franchisee's agreeing to pay a percentage, sometimes as high as six percent, of gross revenues, coupled with his assuming an obligation to provide several channels in the cable system, without charge, to various municipal bodies such as the schools and the fire department. These exactions both reduce the value of the franchise to the cable firm and raise the price to the subscriber above what it would be if the service were an unregulated monopoly.<sup>41</sup>

Thus neither the subscribers nor the cable companies are clear gainers from the current regulatory policies. But they do generate municipal revenues. A tax is imposed on cable subscribers for the benefit of whoever watches the dedicated channels or partakes (either in reduced other taxes or greater municipal services) of the revenues generated by the franchise. In the latter case, indeed, internal subsidization has become conventional taxation in all but name.<sup>42</sup>

Possibly these burdens have been placed on the cable television industry at the behest of competitors, such as theater owners or local television stations (a hypothesis more plausible in the case of the percentage-of-receipts fee than in the case of the dedicated

---

<sup>40</sup> The discussion that follows draws heavily on Posner [45].

<sup>41</sup> See *id.* at pp. 16–19.

<sup>42</sup> Another interesting example of the interplay of explicit taxation and internal subsidies is found in *Student Educational Group Fares* [62], a decision of the Civil Aeronautics Board dealing with a recent Hawaiian statute that granted a tax rebate to airlines establishing a special group rate for students.

channels). In that event, the analogy would be to a tariff—another exercise of the public finance power.

**5. Some general characteristics of regulation and the regulated industries that a public finance approach illuminates**

■ We have seen that viewing regulation as a method of taxation or public finance appears to account, better than alternative views, for a major phenomenon found in regulated industries—internal subsidization. But beyond that, the view provides a consistent explanation of many other features of regulation and regulated industries, some of which fit poorly the received views.

□ **Regulatory control over entry.** Control of entry is an essential feature of regulation under the view advanced here because the adoption of a system of internal subsidies creates false pricing signals. Prices in certain markets must exceed costs if the losses sustained in providing the subsidized services are to be recouped. The price-cost spread in the subsidizing markets will naturally attract new entrants. But their costs may actually be higher than those of the existing firms, in which event their entry would produce a misallocation of resources. Entry would also impair or destroy the system of internal subsidies. With free entry, then, both efficiency and the subsidy scheme would be gravely endangered, so the regulatory agency must control entry.

To be sure, were regulation imposed for the sole benefit of the firms regulated, control over entry would also be necessary to prevent the dissipation through competition of the advantages secured to the incumbent firms by regulation. But not all important instances of entry control can be explained on this ground. The Post Office is not a profit-maximizing enterprise—it is in fact run at a deficit—but new entry into postal service is, and must be, barred, in order to preserve the uniform-rate structure and interclass subsidies that are a prominent feature of the Post Office's operation. Given the financial position of the railroads, it is doubtful that the control of entry in that industry is to be explained in terms of the interests of the regulated firms either.

The theory that regulation seeks to approximate the results of competition cannot explain the control of entry at all. If the regulated firm is constrained to sell at a price approximating cost, there will be no incentive for an inefficient firm to enter. If, despite regulation, the firm is charging a higher price, the matter is more complex. In general, however, assuming that differential pricing is feasible, as seems generally the case in decreasing-cost industries, a new entrant will not be attracted into such an industry by monopoly profits unless it is more efficient than the existing seller. The latter can repel entry by fixing a price near marginal cost to any customer solicited by the new entrant and will, because such a policy will not reduce his profits on any other sales (we have assumed he can maintain different prices) and the alternative is to lose a customer whom it is still profitable to serve even at the reduced price. Unless it is a more efficient firm, the new entrant will have higher costs and will not be able to meet the low price. Thus, in the case where public utility regulation is most plausibly explained in terms of an efficiency rationale—where the industry regulated is a decreasing-cost industry—the rationale still will not explain an important feature of

that regulation, the control of entry, because there is no reason to anticipate inefficient or excessive entry in the absence of public control.

□ **Regulatory review of new construction.** Firms subject to public utility or common carrier regulation are commonly required to obtain the permission of the regulatory agency for any major new construction.<sup>43</sup> This control is to be distinguished from control over entry: it applies whether the purpose of the construction is to enable the firm to enter a new market or to serve an existing market. In arguing that regulation is for the exclusive benefit of the regulated firms, one could point out that such control enables an agency to prevent the firms from expanding production in a way that might undermine cartel pricing. Although the provisions of regulatory statutes giving the agencies authority to fix prices and levels of service, requiring the regulated firms to embody all offerings in published tariffs, and forbidding under severe penalty any deviation from the tariff filings, already give an agency broad authority to enforce cartel pricing,<sup>44</sup> control of new construction makes the agency an even more effective enforcer of the cartel.

A consumer-interest view of regulation also provides an explanation, although not a very satisfactory one, of the control of new construction. While there is little solid basis for fear that an unregulated firm, even if a monopolist, would adopt an extravagant construction program,<sup>45</sup> there are, as noted earlier, some reasons for concern that regulated firms would not minimize costs. But it is unlikely that recognition of the side effects of regulation provides a general explanation of the power over new construction. Many regulated firms subject to the power are not monopolists, and the earlier analysis would not apply. If simply a fear of poor management was in the minds of those who framed the various public utility and common carrier statutes, one wonders why such statutes do not give the regulators more direct authority over management.

I suspect that the framers may have been motivated by a somewhat different concern from those previously mentioned—one that arises from the public finance function of regulation. An illustration, again drawn from the international communications industry, will help explain. In 1967, AT&T, acting this time in concert with the record carriers (and several foreign carriers), applied to the FCC for permission to build a fifth voice-grade cable across the Atlantic Ocean (“TAT 5”).<sup>46</sup> The cost of constructing the facility was estimated to be \$70 million and the planned capacity was 720 voice circuits. Comsat opposed the application. It pointed out that by the time TAT 5 was installed, very large satellites (5000 circuits each) would be in service above the Atlantic and these satellites would provide sufficient capacity to meet all reasonably foreseeable increments of demand for transatlantic telecommunications service at

---

<sup>43</sup> See, for example, the Communications Act of 1934 [78], sec. 214(a); the Natural Gas Act [81]; and the Transportation Act of 1920 [65].

<sup>44</sup> MacAvoy [36] is a case study of the role of regulatory controls in effectuating cartel pricing.

<sup>45</sup> See Posner [47], pp. 573–77.

<sup>46</sup> The incident is discussed by the staff of the President’s Task Force on Communications Policy in [49], pp. 35–49.

a cost per circuit that would be only a small fraction of TAT 5's. AT&T, in reply, noted that the satellites in question might not be in service in time to avoid a shortage. But in that event, judging from subsequent filings and analysis in the proceeding, the economical solution would be to permit queuing, or use peak-load pricing, or launch an additional satellite of an older model. The staff of President Johnson's Task Force on Communications Policy, which analyzed AT&T's application in some depth, concluded that TAT 5 was the least economical alternative.<sup>47</sup>

The cost questions were in fact quite complex and the correctness of the staff's analysis perhaps debatable. The opinions in the case suggest, however, that the FCC itself doubted whether TAT 5 was cost justified. The Commission expressly refused to compare the costs of the cable with those of alternative satellite facilities,<sup>48</sup> adhering to this position in the face of a strong dissenting opinion in which it was urged that the cable was indeed more costly.<sup>49</sup> The majority cannot have been optimistic as to what an analysis of costs would have shown.

In approving the application, the Commission appears to have been strongly influenced by considerations that cannot be understood save in terms of a public finance approach to regulation, such as AT&T's representation that if TAT 5 were approved, it would be able to reduce its transatlantic telephone rates by 27 percent. It is at first glance surprising that the FCC should have been impressed by this offer. If satellites were a cheaper means of meeting demand than the cable, then rates could be reduced by even more than 27 percent if AT&T, rather than building a new cable, leased circuits from Comsat: so why did the Commission refuse to compare cable and alternative satellite costs? The probable explanation lies in Comsat's rate structure. Comsat is wedded, largely it appears for reasons of foreign relations, to a system of uniform global pricing under which the price of a circuit in a Comsat Pacific satellite is roughly the same as the price of a circuit in one of its Atlantic satellites. Because the Atlantic routes are busier, the Atlantic satellites are more fully utilized and the cost per circuit accordingly lower than in the Pacific. But this cost difference is not reflected in the rates, which, as noted, are roughly the same in both markets. Consequently, when AT&T leases circuits from Comsat for transatlantic service, it is forced to pay a considerable premium above the actual cost of the circuits to Comsat, so much so that the price to AT&T (after correcting for certain quality differences) is not clearly lower than the cost to it of circuits in a new cable. It is thus understandable why AT&T should have pushed for approval of TAT 5. But while from its standpoint cable costs may not have been higher than satellite costs, from the broader social standpoint they were (assuming that the staff analysis referred to earlier was correct). It is to prevent unwarranted

---

<sup>47</sup> See *id.*, pp. 36-41a.

<sup>48</sup> Re American Telephone & Telegraph Company [70], pp. 242-43 and n. 4.

<sup>49</sup> Re American Telephone & Telegraph Company [69], pp. 962-63 and [70], pp. 261-62.

<sup>50</sup> The staff of the President's Task Force on Communications Policy estimated that if the price per circuit in an Atlantic satellite had been determined on the basis of the costs of that satellite (and its associated ground facilities), it would have been \$22,400 per half circuit per year in 1970, rather than the actual price of \$31,300 set by Comsat.

investments based on divergences between private and social cost calculations caused by internal subsidization that regulatory agencies must have authority over the construction programs of regulated firms even when entry into a new market is not contemplated.

In this case, to be sure, the agency's exercise of its duty was perfunctory at best. Given the circumstances, however, that is not surprising. First of all, the program of internal subsidies that was jeopardized by the grant of the application—Comsat's policy of uniform global pricing of satellite circuits—is not one that the Commission has particularly encouraged. The motives behind it are rooted, as mentioned, in foreign-policy considerations that are the responsibility of other agencies. By granting the application the Commission was able to obtain immediate rate reductions for one of its constituencies, users of transatlantic telecommunications service, and the costs to the Commission cannot have seemed large.

Secondly, the Department of Defense made strong representations to the Commission<sup>51</sup> that the construction of TAT 5 would promote national security—another example of internal subsidization at work. The Department could have requested an appropriation from Congress to contribute to the cost of building a cable not justified by purely civilian demands; prevailing upon the FCC to authorize such a facility was an alternative mode of financing this defense project. The method of obtaining the “appropriation” and the class of “taxpayers” were different, but the object was the same. The subsidization of defense needs appears to be a rather general feature of regulation.<sup>52</sup>

□ **The duty of the regulated firm to serve and regulatory power over the abandonment of service.** Two long-established and complementary features of the regulatory process are the duty of regulated firms to serve all who demand service and the prohibition against such firms' discontinuing a service without the authorization of the regulatory agency. Although the prohibition of arbitrary refusals of service lies close to the heart of the traditional common-law concept of a public utility or common carrier (as the very name, common carrier, suggests) and is a settled feature of regulatory law,<sup>53</sup> it is difficult to explain under existing views of the purpose in fact of regulation. It is not apparent why regulated firms would want to be placed under such a duty or how they might benefit from it; it is only a little less difficult to see why, from the standpoint of consumer interests, the imposition of such a duty would be thought an appropriate part of the regulatory system. To be sure, a monopolist, if he has his way, will establish a schedule of prices under which fewer customers are served than if a competitive price were set; but once the schedule is adopted there is no reason to expect him to refuse service, on any

---

<sup>51</sup> Alluded to in [69], p. 961, the dissenting opinion of Commissioner Johnson.

<sup>52</sup> See, for example, Bonbright [3], p. 113, and National Transportation Policy in [64].

<sup>53</sup> A typical statement of the duty appears in the Interstate Commerce Act [63]: “It shall be the duty of every common carrier subject to this part to provide and furnish transportation upon reasonable request therefor . . . .” Other examples are given in Jones [28], pp. 376–85. See also *id.* at pp. 26–27. The subject is treated exhaustively by Wyman in [84].

but good business grounds, to any customer willing to pay the price.<sup>54</sup> Arbitrary refusals do not make good business sense. There is similarly no reason to expect a commercial enterprise to abandon a profitable service. Yet regulated firms are forbidden to abandon any route without obtaining the permission of the regulatory agency,<sup>55</sup> and bitterly contested abandonment proceedings are a commonplace occurrence, especially in the railroad industry.

Perhaps these controls are designed in many instances to reinforce regulatory control over the profits of the regulated firms: a firm might refuse or terminate service in order to coerce a higher rate from the customer, or as part of a scheme for enhancing its profits by reducing the level of service on which the rates it was permitted to charge had been based. Possibly they were intended to reinforce ordinary contractual remedies for nonperformance of services considered "essential." But these considerations do not provide a complete explanation. They do not explain why regulatory agencies are empowered to require the extension of utility services to new areas and to prevent the discontinuance of manifestly unremunerative services such as long-distance passenger transportation in the railroad industry. These cases can only be explained, I believe, in terms of a public finance view of regulation. Regulated firms, were they not subject to the duty to institute and not to terminate service, could not be relied upon to implement policies of internal subsidization. For reasons to be noted later, they might still offer some unremunerative services and they might still not always discontinue services when they ceased to be remunerative. But there would be no assurance of their cooperation.

□ **Competitive market structures.** The public finance hypothesis also illuminates some of the important characteristics of the regulated industries themselves (as distinct from characteristics of the regulatory schemes). It suggests, for example, why so many regulated industries do not have a monopolistic structure. A program of internal subsidies does not depend on the regulated industry's being a monopoly. So long as the demand for the industry's product is not perfectly elastic, and so long as the obligation to provide internal subsidies is imposed on all the firms in the industry, such a program is feasible just as it is feasible to impose an excise tax on a competitive industry. It is therefore not surprising, under the view advanced here, that many regulated industries are not monopolistic in structure. To be sure, were regulation imposed solely at the behest of regulated firms, one would also expect many regulated industries to be competitive in structure. But one would not expect so many regulated markets (especially in the communications, power, and water-supply industries) to exhibit pronounced characteristics of natural monopoly. These are the least likely settings for firms to seek governmental protection from competition since the existence of a natural monopoly substantially reduces the danger of competition.

---

<sup>54</sup> See Posner [47], p. 584. I except refusals to serve based upon widespread racial prejudice in a community and refusals to serve business customers for monopolistic reasons. These are special cases, and the scope of the duty to serve is considerably broader.

<sup>55</sup> See Jones [28], pp. 385–95.

□ **Regulated industries produce services.** It is a curiosity that public utility and common carrier industries invariably provide services (in the sense of a good that is difficult or impossible to store or transfer) rather than commodities. The public finance view supplies an answer. A subsidized good or service will not in fact be used by those for whom it is intended if they are free to resell it on the free market, which is why direct subsidies are commonly of services rather than commodities.<sup>56</sup>

□ **Regulated industries provide “infrastructure” services.** The specific complex of controls over entry and over the level and structure of rates that is characteristic of public utility and common carrier regulation is confined, for the most part, to the transportation, telecommunication, and power (electricity and gas) industries. Neither of the received views of regulation explains adequately why these particular industries have been singled out. The consumer-interest view of regulation would suggest that these were industries in which monopolistic misallocations of resources were most likely to occur, yet, as mentioned, many of the industries are naturally competitive rather than monopolistic, while a number of important industries, such as computers, drugs, newspapers, and certain non-ferrous metals, which appear to have monopoly problems, sometimes quite serious ones, have escaped regulation. On the other hand, if we assume that regulation is imposed primarily for the benefit of the regulated firms, it must be shown why other industries have not obtained the same kind of regulation as public utilities and common carriers.

A partial explanation of the identity of the regulated industries may be that society frequently subjects to the public utility type of control services that it wants provided on the broadest possible basis (in a sense to be defined). The regulated industries are part of the “infrastructure” of economic growth. Adequate transportation, communications, and power (especially electrical) must be in place before the development of modern industry is possible, and most countries, including this one at various periods, have undertaken to subsidize these services or provide them directly in the hope thereby of attracting industrial developers.<sup>57</sup> One can deny the necessity or appropriateness of this state promotional role but hardly its prevalence. And internal subsidization is one method whereby the expansion of the infrastructure services can be promoted.

To be sure, it is not “expansion” in any simple or obvious sense that is involved. In the case of a naturally competitive industry, internal subsidies expand the provision of service to one class of customers, the beneficiaries of the program, but contract it to another: those who must pay a higher price to defray the subsidy and who consequently demand (and are supplied) a smaller quantity. The overall output of the industry is not necessarily larger, and may (as we saw in discussing value-of-service pricing) be smaller. If the industry is monopolistic in structure and it is not feasible to control its monopoly profits directly, a program of internal subsidies may well

---

<sup>56</sup> Shoup [55], p. 160; Stigler [58], p. 5.

<sup>57</sup> See, for example, Locklin [34], pp. 101–11, and Smead [56]. The theory of social overhead capital (as investment in infrastructure is often called by economists) is discussed by Hirschman [25], pp. 83–97.

bring about a larger output than otherwise. But in either case it would appear that the primary effect of such a program is not to increase the amount of transportation, communication, or power produced but rather to extend the service to classes of customers and geographical areas that might not be served in a free market.

Such a result is nonetheless consistent with the thinking that underlies the desire to force the creation of an adequate infrastructure rather than let the market take its course. The basic assumption, correct or incorrect, is that private enterprise, due to lack of foresight, or imperfections in the capital market, or external economies, will forgo many investments in infrastructure that would be socially profitable.<sup>58</sup> One can argue from this that it is the role of the state to encourage precisely those infrastructure services that are unremunerative.

This view may be reinforced in some cases by another: concern with geographical concentration of population and economic activity. A program of internal subsidies that denies the cost advantages of proximity and density, as is often the case, encourages greater geographic dispersion. Cost advantages based on location, it need hardly be said, are no less real than those based on other factors. But governments, including our own, have frequently followed policies aimed at denying those advantages. Utility regulation is perhaps one of them.

The industries in which we find internal subsidies are commonly also recipients of at least some direct subsidies.<sup>59</sup> This correlation supports the view of regulation as a method of public finance, especially where, as in the case of the electrical and telephone subsidies doled out by the Rural Electrification Administration, the recipients of direct subsidies are not members of the industry at all (in the REA case, they are consumer cooperatives). In such a case the established firms in the industry benefit only insofar as the existence of the direct subsidy reduces the pressure on them to provide an internal subsidy, and the subsidy scheme is more convincingly interpreted as a method of obtaining greater service than as a device for enriching corporate treasuries.

To suggest that regulation is a method of promoting the expansion of infrastructure services is not, of course, to explain why it is chosen in preference to alternative methods, such as direct subsidies, or why, with respect to some infrastructure services, such as education, the public utility approach plays a very subordinate role.<sup>60</sup> A framework for answering this question is sketched in part 6, where we look at some of the advantages and disadvantages of regulation in comparison with other methods of taxation, and a highly tentative answer is suggested.

The infrastructure explanation for the identity of the regulated industries is far from being completely satisfactory. It hardly seems applicable when an internal subsidy is used to retard the decline of an old industry, such as railroad passenger service or telegraph service. In addition, the economic case for subsidizing infrastructure

---

<sup>58</sup> See, for example, Hagen [19], pp. 126–29, and Kindleberger [32].

<sup>59</sup> Cf. Harriss [21], p. 270.

<sup>60</sup> Although not an entirely negligible one, as attested by free cable-television channels for schools, the public-affairs programming obligations of broadcast licensees, and the proposed free interconnection for educational television.

services is often dubious. And internal subsidization seems a somewhat curious way to encourage the expansion of an industry since, as mentioned, the cost of the subsidy is borne by customers of the industry. Indeed, the obligation to provide service to all at a uniform price may retard the undertaking of new extensions of service.<sup>61</sup>

At the least, these considerations suggest that a thoroughgoing justification of internal subsidies on efficiency grounds is impossible. One can easily find examples where an internal subsidy works directly contrary to the dictates of efficient resource allocation. Thus, the subsidization of commuter railroad service aggravates an existing imbalance between private and social costs caused by the fact that individuals who are employed in cities and utilize urban public services can escape the costs of those services by living in a suburb and commuting. It would appear, therefore, that internal subsidies are frequently designed to redistribute wealth rather than to correct imperfections in the market.

■ **Limitations of the device.** To summarize the discussion at this point, there is persuasive evidence that an important purpose in fact of public utility and common carrier regulation is to compel, by the device of the internal subsidy, the provision of certain services in quantities and at prices that a free market would not offer, much as the conventional taxing-spending power is used to the same end. Serious discussion of the public finance component of regulation has been retarded, however, by a tendency to dismiss it out of hand as an implausible and inappropriate alternative to more conventional exertions of the taxing power. Two objections are usually advanced as conclusive. The first is that internal subsidization distorts the efficient allocation of resources; the second, that it tends to be arbitrary and inequitable. One sometimes hears it said, too, that taxation is the proper business of the legislature and not of regulatory agencies.

1) *Delegation.* To take the last point first, it is difficult to understand why the delegation of a part of the taxing power to appointive agencies, the regulatory commissions, should be thought to offend the principles on which our government is organized. Congress, acting from imperative reasons of practicality, has delegated much of its lawmaking power to appointive agencies. The Federal courts provide a conspicuous example, and the Internal Revenue Service one that is directly in point.

2) *Efficiency.* It is true that internal subsidization, by forcing prices in some markets above cost and prices in others below, distorts the allocation of resources. It creates a secondary inefficiency as well: the entry of new competitors into the high-price markets must be prevented by the regulatory agency lest the source of the internal subsidy be wiped out.<sup>62</sup> Where the high-price market is

---

<sup>61</sup> Coase [6], p. 139. The effect of uniform price systems with which Professor Coase was concerned—denial of service to high-cost customers willing to pay the full cost of serving them—can be avoided by combining a uniform price system for most customers with a system of surcharges for those who would not be served at the uniform price. This appears on casual observation to be the practice of the telephone industry in this country.

<sup>62</sup> For a good example, see Caves [5], p. 314.

## 6. Internal subsidization compared with other methods of public finance

a natural monopoly, this is not an acute problem, but of course not all markets subject to regulation are naturally monopolistic.

The criticism of internal subsidization as inefficient points to a real characteristic of the device; but as a criticism it is superficial. It measures the device against an ideal standard, and of course finds it wanting. The proper comparison is to other exercises of the taxing power. All methods of taxation distort the “optimum” allocation of resources—optimum, that is, without regard to any need or demand to provide certain services publicly—and there are no *a priori* grounds for assuming that excise taxes, such as the internal-subsidy programs imposed by regulatory agencies, produce worse misallocations than income or other taxes.<sup>63</sup> To consider an important example, the exemption from income taxation of the real but not pecuniary income generated by housewives must cause a significant misallocation of resources by inducing many women to stay at home who would be more productive in other employments. The administrative costs of implementing a broader income concept would be so great, however, that this exemption is probably a permanent feature of income taxation. Because of pervasive and ineradicable distortions of this kind, it is not obvious that raising income tax rates would be a more efficient method of providing particular services at below-market prices than internal subsidization.<sup>64</sup> Indeed, insofar as the burden of internal subsidies tends to be borne by customers whose demands are highly inelastic, the allocative effects may be less adverse than those of alternative taxation methods.<sup>65</sup> And in those cases where the regulated firms are obtaining monopoly profits, the adverse allocative effects of the tax will be even fewer.<sup>66</sup>

Internal subsidies are also criticized on the ground that a subsidy in kind is inefficient compared to an unrestricted cash subsidy, because different people have different needs and wants. This is a valid and important point but it is not a criticism of internal subsidies as such, since it applies with equal force to most direct subsidies.

3) *Equity*. Because the determination of the incidence of particular taxes is immensely complex, it is very difficult to gauge the effect of internal subsidies on the distribution of income. At a rough guess, internal subsidization may sometimes benefit the poor<sup>67</sup> but has no general tendency to do so; and as our commuter example shows it may sometimes work in the opposite direction. But poverty is not the only possible justification for the redistribution of income. It is notable that internal subsidization is frequently employed to bolster declining services or sectors; perhaps in these cases it is felt that there are important reliance interests (for example, in location proximate to a railroad line) that deserve protection. And even if no consistent equity justification is possible, that is no special criticism

---

<sup>63</sup> See Friedman [16], pp. 56–67, and Little [33], p. 608.

<sup>64</sup> Cf. Harberger [20], p. 58.

<sup>65</sup> Cf. Musgrave [39], p. 157.

<sup>66</sup> Cf. Shilling [54], p. 224; Shoup [55], p. 276; Dirlam and Kahn [13], p. 494.

<sup>67</sup> Some rather dubious examples are given in Bonbright [3], pp. 111–12. Lower prices to the poor could in some cases be explained as profit-maximizing price discrimination, and his examples may well be of this type. A better example is provided by the low rail rates for agricultural commodities (see p. 26 and fn. 21 *supra*), but it is possible that the benefits are largely captured by farmers (who may or may not be poor) rather than by consumers. Staudinger [57], p. 259, argues for using public utility pricing to redistribute income to the poor,

of internal subsidies: the redistributive effects of tax-cum-direct-subsidy programs appear in a surprising range of cases to be perverse.<sup>68</sup> If one is to oppose internal subsidies on equity grounds, it must be as part of a broader objection to the redistributive policies of the state.

I turn now to some other, less frequently discussed attributes of regulation as a method of public finance.

4) *Enforcement.* An important characteristic of taxation by regulation is difficulty (and expense) of enforcement. A firm that finds the provision of an unremunerative service irksome may try to terminate it by drastically reducing the quality of the service and then citing the resulting fall in demand as evidence that the public no longer wants the service. This is not so transparent a gambit as it may seem. Since the public is not paying the full cost of the service, it has a natural tendency to demand a very high (and correspondingly costly) level of service. The specification of an appropriate level involves an essentially arbitrary judgment and accordingly gives the firm some room for maneuver. Evidently degradation of service has played an important role in the termination of railroad passenger operations.<sup>69</sup>

The tendency of regulated firms to cheat in providing unremunerative services is probably quite general since, unless regulation is more effective than anyone thinks, a penny saved in skimping on an unremunerative service will not result immediately in a full penny reduction in the rates paid by customers of the firm's lucrative services. The finding in a recent study that the rates set by publicly owned electric utilities<sup>70</sup> are more uniform than those set by privately owned electric utilities supports this suggestion. Uniform rates, we saw, are a common method of internal subsidization; and one would expect a privately owned company to resist providing unremunerative services more energetically than a publicly owned one.

The tendency to cheat is not entirely a bad thing. It may result in a smaller subsidy than if direct subsidization were used, and given the forceful objections to many subsidies, this may be distinctly to the good.

5) *Public scrutiny.* A troubling characteristic of the internal subsidy is its low visibility, which impedes responsible review. The amounts and recipients of direct subsidies are ordinarily specifically stated, but this is not the case with internal subsidies. Since information is not a free good, a subsidy program whose magnitude requires computation is less apt to be challenged than one whose magnitude is patent.

This is a general criticism of hidden subsidies, of which internal subsidies in the regulated industries are only one variety. And it is easily overstated: extravagant subsidy programs sail through Congress with monotonous regularity. Full disclosure is a far from dependable test of whether legislation in the public interest will be adopted, because the public does not vote on specific pieces of legislation, but on representatives, and it is demonstrable that in a repre-

---

<sup>68</sup> Stigler [58], p. 1.

<sup>69</sup> Doubtless encouraged by the ICC's recent holding that it has no authority over the quality of rail passenger service [76].

<sup>70</sup> Peltzman [44].

sentative system much legislation benefiting special interests at the expense of the larger public will be enacted.<sup>71</sup> This is a basic insight of the effective-political-group theory of regulation. Furthermore, given the size of the Federal budget, the disclosure in an appropriation hearing of the amount of a subsidy may not always be an effective method of assuring a responsible review of the proposal's merits. Suppose that buried in the Defense Department's appropriation request there had been a small item for a contribution to the cost of building a transatlantic cable: can one be confident that it would have received the careful scrutiny of Congress?

Despite the last point, the concern about adequacy of scrutiny has greatest force, I believe, precisely with regard to internal subsidies for national defense. The Defense Department's role in the TAT 5 matter affords a good illustration. Had the Department been forced to include the item in its budgetary request to Congress, it would have had to weigh its importance against that of other national-defense programs. The defense budget is not limitless. The inclusion of the cable item might have compelled the Department to modify some other request. In the context of a regulatory proceeding, however, the cable represented a free good to the Department. The Department had no incentive to evaluate the benefits of TAT 5 to the national defense objectively; indeed, it had an incentive to exaggerate those benefits. The FCC could not exercise a critical scrutiny because it has no competence to deal with military questions. The competent agencies—Congress and the Bureau of the Budget (which reviews all Federal budgetary proposals before submission to Congress)—were bypassed.

6) *Manageability of regulation.* Another problem with internal subsidization is that it complicates an already barely manageable regulatory process. Because there is no objective basis for balancing off distributive benefits against allocative costs,<sup>72</sup> an agency concerned with subsidizing worthy groups is deprived of a clear-cut standard for resolving controversies over pricing and entry. Clear and definite standards are necessary to tolerable regulation.<sup>73</sup> Without a definite standard at the agency level, moreover, judicial review, a potentially

---

<sup>71</sup> See Stigler [59].

<sup>72</sup> In his recent study of the Federal Power Commission [35], pp. 288–89, Paul MacAvoy makes a valiant but, I believe, unsuccessful effort to do so. His position on determining the dollar value of an income redistribution is “that the government should decide, and it indicates value by the amount that consumer *X* can keep after taxes . . . .” The difficulty with such an approach is that the scheme of Federal taxation does not reflect any consistent or intelligible policy toward the equity of income redistribution—unless we are to assume, for example, that because some dividend income is exempt from tax, stockholders are to be deemed a favored class for purposes of evaluating the redistributions brought about by a monopolist or a regulatory agency. See Pechman [43] for a lively recent discussion. And many of the apparently distributive features of Federal income taxation, including the progressive principle itself, have been justified on grounds (such as benefits received) that have nothing to do with the equity of redistributing income, while other features (such as the nontaxability of real income received in the form of reductions in the prices of consumer goods) reflect purely administrative considerations.

<sup>73</sup> Friendly [17]. For a somewhat different path to the same conclusion, see Posner [46], pp. 84–85.

important check on regulatory excesses,<sup>74</sup> is likely to be ineffectual; the agency can give a plausible justification for any result. Multiple and conflicting standards also breed corruption.<sup>75</sup>

7) *Private demand.* Taxation by regulation, to be feasible, requires that there be sufficient demand in the private market to justify the imposition of the burden of the subsidy on the regulated firms. Where there is not, as in the railroad industry, the results can be disastrous for the industry. One may hazard the guess that regulation has frequently been the principal means of subsidizing infrastructure services for which there is a strong private demand, while in areas like national defense and education, where the market demand is probably small in relation to the amount of service that the state wishes to provide, other methods of subsidization have predominated.

□ **And some advantages.** The balance of advantages is not wholly against the choice of the internal subsidy as a method of public finance. We have indicated several respects in which it may be preferable to other methods, and there are others.

1) *Administrative expense.* Although enforcement of internal subsidization can, as mentioned, be quite costly (railroad abandonment proceedings are a case in point), there are certain offsetting factors. Since no cash transfers are involved in internal subsidization, it is possible to dispense with the frequently elaborate apparatus of a formal transfer program—application forms, disbursement machinery, and the like. Often, too, a program of internal subsidies is implemented simply by the regulated firms' averaging the costs of many customers in setting a rate, and where this is done the firms avoid the expenses that would be incurred in identifying the costs of finer groups of customers and adopting a more complex rate structure tailored to the different costs. This is not to suggest that expenses incurred in implementing the price system are normally wasted; but once it is decided not to allow the price system to control the allocation of resources, a choice implicit in the decision to subsidize, the elimination of some of those expenses may represent a real saving.

2) *Legislative capacity.* By shifting taxing power from Congress (or state legislative bodies) to administrative agencies, internal subsidization economizes on the legislature's time. This is an especially important consideration where the subsidy is of a kind that requires frequent adjustment or review. The ability of a legislature to transact business is obviously limited. Among the ways in which it can be conserved, perhaps the delegation of minor taxing functions to regulatory agencies is relatively efficient.

3) *Protection of expectations.* At least when imposed on a service from the outset, internal subsidies may be less disruptive of public and

---

<sup>74</sup> An expanded role for the courts in the review of regulatory action is urged in Turner [61], p. 386. His position can be defended on the ground that judges are more insulated from political-group pressures than regulatory agencies. Cf. Posner [46], p. 89.

<sup>75</sup> The classic instance is the corruption that beset the FCC, at the highest levels, in the 1950's. The problem revolved around the initial grant of broadcast licenses, where the Commission applied no standard but used a check-list of criteria, enabling any preconceived result to be rationalized. Jones [28], pp. 1081–84, lists 15 criteria.

commercial expectations than other new taxes. An example will illustrate. Suppose a community has pending before it several applications for a cable television franchise and would like to use a few channels in any cable television system that is constructed for municipal functions such as education. And suppose further that the feasible alternative methods of obtaining this service have been narrowed to two: a tax on the gross receipts of the barbers in the community, the proceeds to be used to purchase the channels from the cable franchisee, and a condition in the franchise requiring the franchisee to provide the channels to the school system at no charge. If the first alternative is chosen, the result will be a rise in the cost (and hence presumably price) of barbering, which will lead to a fall in the amount of barbering demanded and supplied. As a result, some of the resources used in barbering in the community will be idle during the period in which they are being redeployed. And there will be an outcry from the barbers. These economic and political costs, incurred by virtue of the change in the economic conditions of the business brought about by a new tax, can be avoided if the second alternative, an internal subsidy by the cable industry, is selected. Since the costs of the cable system are now higher, a smaller system will be built. But the efficient scale (consistent with the obligation to provide free channels to the franchising authority) will be known in advance; there will be no waste in achieving it, as in the barbering example. In fact one observes that public utility and common carrier regulation has typically been imposed upon new services, where it was possible by a system of internal subsidies to finance desired extensions of the service without disturbing settled activities. And perhaps these considerations explain why municipalities have latched onto cable television as an important new source of revenue.<sup>76</sup>

Nonetheless, the explanation is severely limited. The alternatives in our example were too narrow: the municipality could also have placed a gross-receipts or other tax on cable service and raised the money for the free channels that way. It did not have to use internal subsidization, although we have previously discussed some reasons why internal subsidization might sometimes be preferred to alternative forms of excise taxation.

4) *Justice*. There may be some appeal to the notion that it is more “just” for other customers of the same industry to bear the cost of a subsidy of the industry’s service than to distribute that cost among the taxpaying public at large. The notion is a little peculiar, however. It is one thing to say that those who benefit from a service should bear its costs, and quite another to impute the cost of a subsidy to those customers who are quite prepared to pay the full cost of serving them.

A final reason for the choice of internal subsidization over alternative methods of public finance has nothing to do with its relative merits. The regulated firms may cast their weight on the side of the internal subsidy, viewing customers who enjoy subsidized rates as useful allies in the maintenance of regulatory barriers to entry. Subsidizing some customers may be the “price” that the franchised monopolist pays for his monopoly. Perhaps careful study would disclose that most regulation is demanded by and supplied to a coalition

---

<sup>76</sup> See, for example, the Mayor’s Advisory Committee [41].

of regulated firms and those of their customers who receive services below cost as a consequence of regulation.

■ I trust that the foregoing remarks will not be construed as a “defense” of taxation (and subsidization) by regulation. They may, however, help explain the prevalence and tenacity of the practice, and they do suggest that, short of a thorough overhauling of government subsidy policy, it is less easy to condemn the practice out of hand as inefficient and inequitable than has usually been assumed. Perhaps few subsidies are in the public interest; there may still be cases where, given a decision to subsidize, regulation is the cheapest means of doing so.

But if we are stuck with taxation by regulation, perhaps we are not stuck with its worst features. I propose two modest reforms. The first is that agencies and reviewing courts insist, in proceedings where the maintenance of an internal subsidy is an issue, that the amount and cost of the subsidy, together with the identity of the recipients and of the payors, be calculated and placed in the public record. Perhaps this would eliminate some of the more captious instances of the phenomenon; at least it would bring an important issue of public policy into the open.

Second, more consideration should be given to the most efficient method of attaining the ends of internal subsidization. Accepting the decision to subsidize a specific service and to impose the cost of the subsidy on other customers of the firm providing the service, there may be better ways of achieving this end than control of prices, entry, abandonments, and the like by a regulatory agency. In particular, an explicit excise tax (such as the percentage-of-gross-receipts fee in many cable television franchises), with the proceeds earmarked for the service that the state wants to subsidize, may be preferable to the internal subsidy proper because it entails no limitation on entry into the high-price market; lump-sum fees may be preferable to either.<sup>77</sup> A likely reason why such alternatives are rarely considered is that the usual regulatory agency lacks authority to impose an explicit tax or other fee. In franchise regulation, as the case of cable television suggests, this option is open. Perhaps, therefore, a modest enlargement of the taxing power of regulatory agencies, to permit them to exact a uniform and limited fee from any firm desiring to enter a regulated market in lieu of other regulatory controls, would foster the more efficient use of what appears to be a settled device of public finance.

■ This paper merely scratches the surface of an interesting and important question of public policy. I have tried to show that certain views of the purpose in fact of public utility and common carrier regulation—that it is to approximate competitive results, or that it is to benefit the regulated firms—fail to account for a number of significant observed features of the regulatory process and the regulated industries. And I have argued that a consistent and comprehensive explanation of those features requires that we assign an important place to taxation and subsidization among the purposes that regula-

## 7. Some practical suggestions

## 8. Conclusion

---

<sup>77</sup> See Posner [45] pp. 19–20.

tion in fact serves. I have attempted further to compare taxation by regulation with other methods of taxation and subsidization, in the hope of assisting evaluation of the pros and cons of alternative methods in particular cases. What I have not attempted to do is explain why some groups are subsidized and others not, or why the same group will receive some internal subsidies but not others (in the case of fire departments, for example, free water but not free telephone service). These fascinating and important questions, which require a better understanding of the magnitude and incidence of taxation by regulation than existing information permits, constitute the agenda for further research into a heretofore rather neglected aspect of public regulation.

## References

1. AVERCH, H. and JOHNSON, L. L. "Behavior of the Firm Under Regulatory Constraint," *American Economic Review*, Vol. 52 (December 1962), pp. 1053-69.
2. BARATZ, M. "Cost Behavior and Pricing Policy in the Post Office," *Land Economics*, Vol. 38 (1962).
3. BONBRIGHT, J. C. *Principles of Public Utility Rates*. New York: Columbia University Press, 1961.
4. BOWMAN, W. "Tying Arrangements and the Leverage Problem," *Yale Law Journal*, Vol. 67 (1957).
5. CAVES, R. E. *Air Transport and its Regulators: An Industry Study*. Cambridge, Mass: Harvard University Press (Economic Studies Series, No. 120), 1962.
6. COASE, R. H. "The Economics of Uniform Pricing Systems," *Manchester School of Economic and Social Studies*, Vol. 15 (1947).
7. ———. "The Marginal Cost Controversy," *Economica*, Vol. 13 (1946), pp. 169-82.
8. ———. "The Nationalization of Electricity Supply in Great Britain," *Land Economics*, Vol. 26 (1950).
9. ———. "The Postal Monopoly in Great Britain: An Historical Survey," In *Economic Essays in Commemoration of the Dundee School of Economics, 1931-1955*, J. Eastham, editor, 1955.
10. CONANT, M. *Railroad Mergers and Abandonments*. Berkeley, California: University of California Press, 1965.
11. CREW, M. "Electricity Tariffs," In *Public Enterprise*. Edited by Ralph Turney. 1968.
12. DAVIDSON, R. *Price Discrimination in Selling Gas and Electricity*. 1954.
13. DIRLAM, J. and KAHN, A. "The Merits of Reserving the Cost Savings from Domestic Communications Satellites for the Support of Educational Television," *Yale Law Journal*, Vol. 77 (1968).
14. EADS, G. "The Effect of Regulation on the Cost Performance and Growth Strategies of the Local Service Airlines," unpublished manuscript, Princeton University Department of Economics, 1970.
15. FRIEDLAENDER, A. F. *Dilemma of Freight Transport Regulation*. Washington, D. C.: The Brookings Institution, 1969.
16. FRIEDMAN, M. *Price Theory: A Provisional Text*. Chicago, Ill.: Aldine Press, 1966 (rev. ed).
17. FRIENDLY, H. J. *Federal Administrative Agencies: The Need for Better Definition of Standards*. Cambridge, Mass.: Harvard University Press, 1962.
18. GARFIELD, P. and LOVEJOY, W. *Public Utility Economics*. Englewood Cliffs, New Jersey: Prentice-Hall, 1963.
19. HAGEN, E. E. *The Economics of Development*. Homewood, Ill.: Richard D. Irwin, 1968.
20. HARBERGER, A. C. "Taxation, Resource Allocation, and Welfare," in NATIONAL BUREAU OF ECONOMICS RESEARCH and THE BROOKINGS INSTITUTION, *The Role of Direct and Indirect Taxes in the Federal Revenue System: A Conference Report*. Princeton, N. J.: Princeton University Press, 1964.
21. HARRISS, L. "Subsidies in the United States," *Public Finance*, Vol. 16 (1961).

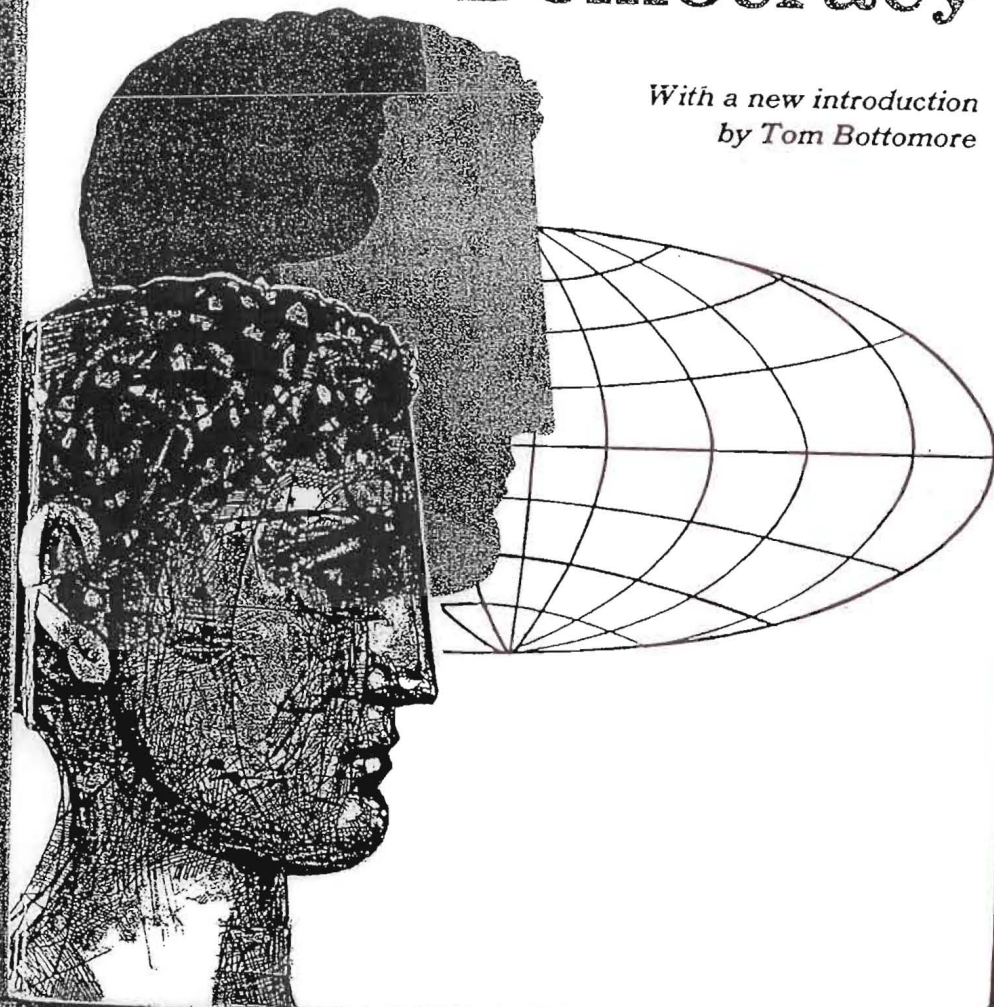
22. HENDERSON, A. "The Pricing of Public Utility Undertakings," *Manchester School of Economic and Social Studies*, Vol. 15 (1947).
23. HILTON, G. W. *The Transportation Act of 1958: A Decade of Experience*. Bloomington, Indiana: Indiana University Press, 1969.
24. HIRSHLEIFER, J., DE HAVEN, J. and MILLIMAN, J. *Water Supply: Economics, Technology, and Policy*. Chicago, Ill.: University of Chicago Press, 1960 (rev. ed., 1969).
25. HIRSCHMAN, A. O. *The Strategy of Economic Development*. New Haven, Conn.: Yale University Press, 1958.
26. ILLINOIS, STATE OF. "An Act to provide for the establishment of ferries, toll bridges, and turnpike roads," *Revised Code of Laws of Illinois*, Sec. 5 (1827).
27. JOHNSON, L. *Communications Satellites and Telephone Rates: Problems of Government Regulation*. Rand Memorandum RM-2845-NASA (October 1961).
28. JONES, W. K. *Cases and Materials on Regulated Industries*. 1967.
29. KENNEDY, J. "Structure and Policy in Postal Rates," *Journal of Political Economy*, Vol. 65 (1957).
30. KEETON, R. E. and O'CONNELL, J. *Basic Protection for the Traffic Victim: A Blueprint for Reforming Automobile Insurance*. Boston, Mass.: Little, Brown & Company, 1966.
31. KEIG, N. G., FRISTOE, C. W. and GODDARD, F. O. "A Critique of the Policy Objectives of Publicly Owned Water Utilities," Unpublished manuscript, University of Florida, 1970.
32. KINDLEBERGER, C. P. *Economic Development*. New York, N. Y.: McGraw Hill, 1958 (2nd ed. 1965).
33. LITTLE, I. M. D. "Direct Versus Indirect Taxes," *Economic Journal*. Vol. 61 (1951). Reprinted in AMERICAN ECONOMICS ASSOCIATION, *Readings in Welfare Economics*, 1969.
34. LOCKLIN, D. P. *Economics of Transportation*. Homewood, Ill.: Richard D. Irwin, 1966 (6th ed.).
35. MACAVOY, P. W. "The Effectiveness of the Federal Power Commission," *Bell Journal of Economics and Management Science*, Vol. 1, No. 2 (Autumn, 1970), pp. 271-303.
36. ———. *The Economic Effects of Regulation: The Trunk-Line Railroad Cartels and the Interstate Commerce Commission before 1900*. Cambridge, Mass.: MIT Press, 1965.
37. MEYER, J. R., KAIN, J. F. and WOHL, M. *The Urban Transportation Problem*. Cambridge, Mass.: Harvard University Press, 1965.
38. ———, PECK, M., STENASON, J. and ZWICK, C. *The Economics of Competition in the Transportation Industries*. Cambridge, Mass.: Harvard University Press, 1959.
39. MUSGRAVE, R. *The Theory of Public Finance: A Study in Public Economy*. New York: McGraw Hill, 1959.
40. NELSON, J. C. *Railroad Transportation and Public Policy*. Washington, D. C.: The Brookings Institution, 1959.
41. NEW YORK MAYOR'S ADVISORY COMMITTEE ON CATV AND TELECOMMUNICATIONS. *A Report on Cable Television and Cable Telecommunications in New York City*. New York, 1968.
42. NOVE, A. "Internal Economies," *Economic Journal*, Vol. 79 (1969).
43. PECHMAN, J. A. "The Rich, the Poor, and the Taxes They Pay," *The Public Interest*, No. 17 (Fall 1969), pp. 21-43.
44. PELTZMAN, S. "Pricing in Public and Private Enterprises: Electric Utilities in the United States," *Journal of Law and Economics*, forthcoming.
45. POSNER, R. A. *Cable Television: The Problem of Local Monopoly*. Rand Memorandum RM-6309-FF (May 1970).
46. ———. "The Federal Trade Commission," *University of Chicago Law Review*, Vol. 37 (1969).
47. ———. "Natural Monopoly and Its Regulation," *Stanford Law Review*, Vol. 21 (1969).
48. PRESIDENT'S TASK FORCE ON COMMUNICATIONS POLICY. *Final Report*. Washington, D. C.: Government Printing Office, 1968.
49. ———. *Organization of the United States International Communications Industry*. U. S. Department of Commerce, Clearinghouse for Federal Scientific and Technical Information, PB 184 424 (June 1969). Staff Paper 2.

50. ———. *The Western Union Telegraph Company*. U. S. Department of Commerce, Clearinghouse for Federal Scientific and Technical Information, PB 184 418 (June 1969). Staff Paper 5.
51. SARGENT, J. R. "Nationalized Industries and Public Policy," in *The Lessons of Public Enterprise*, edited by Michael Shanks (1963).
52. SHARP, C. *Problems of Urban Passenger Transport with Special Reference to Leicester*. 1967.
53. SHEPHERD, W. G. "Cross-Subsidizing and Allocation in Public Firms," *Oxford Economic Papers*, Vol. 16 (N.S., 1964).
54. SHILLING, N. *Excise Taxation of Monopoly*. New York: Columbia University Press, 1969.
55. SHOUP, C. S. *Public Finance*. Chicago, Ill.: Aldine Press, 1969.
56. SMEAD, E. E. *Governmental Promotion and Regulation of Business*. New York: Appleton-Century Crofts, 1969.
57. STAUDINGER, H. "Social Rates in Electricity," *Social Research*, Vol. 3 (1936).
58. STIGLER, G. J. "Director's Law of Public Income Redistribution," *Journal of Law and Economics*, Vol. 13 (1970).
59. ———. "Theory of Economic Regulation," *Bell Journal of Economics and Management Science*, Vol. 2, No. 1 (Spring 1971).
60. ———. *The Theory of Price*. New York: The Macmillan Company, 1966 (3rd ed).
61. TURNER, D. "The Scope of Antitrust and Other Economic Regulatory Policies," *Harvard Law Review*, Vol. 82 (1969).
62. U. S., CIVIL AERONAUTICS BOARD. *Student Educational Group Fares*. Docket No. 22402, order of investigation and suspension, July 29, 1970.
63. U. S. Code. "Interstate Commerce Act," Vol. 49, Sec. 1(4).
64. ———. "National Transportation Policy," Preceding Vol. 49, Sec. 1.
65. ———. "Transportation Act of 1920," Vol. 49, Secs. 1(18), 1(20), and 1(22).
66. U. S., CONGRESS, HOUSE, COMMITTEE ON INTERSTATE AND FOREIGN COMMERCE. *Transportation Acts Amendments, 1962, Hearings on*. 87th Cong., 2nd Session.
67. U. S. DEPARTMENT OF TRANSPORTATION. *Price Variability in the Automobile Insurance Market*. Automobile and Compensation Study, August 1970.
68. U. S., FEDERAL COMMUNICATIONS COMMISSION. *Re American Telephone and Telegraph Company*, 37 F.C.C. 1151 (1964).
69. ———. *Re American Telephone and Telegraph Company*, 11 F.C.C. 2d 957 (1968).
70. ———. *Re American Telephone and Telegraph Company*, 13 F.C.C. 2d 235 (1968).
71. ———. *Annual Report, 1969*. Washington, D. C.: U. S. Government Printing Office, 1970.
72. ———. *Authorized Entities and Authorized Users Under the Communications Satellite Act of 1962*, 4 F.C.C. 2d 421 (1966).
73. ———. *Interconnection Service*, 14 F.C.C. 2d 599 (1968).
74. ———. *Public Service Responsibility of Broadcast Licensees* (1946).
75. ———. *Re Western Union Telegraph Company*, 12 F.C.C. 2d 980 (1968).
76. U. S., INTERSTATE COMMERCE COMMISSION. *Adequacy—Passenger Service—Southern Pac. Co. Between California and Louisiana*, 335 I.C.C. 415 (1969).
77. ———. *Increased Freight Rates, E.W. & S. Territories, 1956*, 299 I.C.C. 429, 441, 451–59 (1956).
78. *U. S. Statutes at Large*. Communications Act of 1934, as amended by the Public Broadcasting Act of 1967 (U. S. Statutes at Large, vol. 81, p. 365), U. S. Code Vol. 47, Secs. 1 *et seq.*
79. ———. Communications Satellite Act of 1962, Vol. 76; U. S. Code, Vol. 47, Secs. 701 *et seq.*
80. ———. Hoch-Smith Resolution, Vol. 43 (1925).
81. ———. Natural Gas Act, sec. 7(c), as amended. U. S. Code, Vol. 15, sec 717 ff.
82. WATKINS, G. P. "Street-Railway Rates, with Especial Reference to Differentiation," *Quarterly Journal of Economics*, Vol. 25 (1911).
83. WILSON, G. "Effects of Value-of-Service Pricing Upon Motor Common Carriers," *Journal of Political Economy*. Vol. 63 (1965).
84. WYMAN, B. *Public Service Corporations*. Vol. 1 (1911), Parts II–IV.

*Joseph A. Schumpeter*

**Capitalism,  
Socialism  
and  
Democracy**

*With a new introduction  
by Tom Bottomore*



the equilibrium in perfect competition, and still harder to preserve, but the "beneficial" competition of the classic type seems likely to be replaced by "predatory" or "cutthroat" competition or simply by struggles for control in the financial sphere. These things are so many sources of social waste, and there are many others such as the costs of advertising campaigns, the suppression of new methods of production (buying up of patents in order not to use them) and so on. And most important of all: under the conditions envisaged, equilibrium, even if eventually attained by an extremely costly method, no longer guarantees either full employment or maximum output in the sense of the theory of perfect competition. It may exist without full employment; it is *bound* to exist, so it seems, at a level of output below that maximum mark, because profit-conserving strategy, impossible in conditions of perfect competition, now not only becomes possible but imposes itself.

Well, does not this bear out what the man in the street (unless a businessman himself) always thought on the subject of private business? Has not modern analysis completely refuted the classical doctrine and justified the popular view? Is it not quite true after all, that there is little parallelism between producing for profit and producing for the consumer and that private enterprise is little more than a device to curtail production in order to extort profits which then are correctly described as tolls and ransoms?

## CHAPTER VII

## THE PROCESS OF CREATIVE DESTRUCTION

THE theories of monopolistic and oligopolistic competition and their popular variants may in two ways be made to serve the view that capitalist reality is unfavorable to maximum performance in production. One may hold that it always has been so and that all along output has been expanding in spite of the secular sabotage perpetrated by the managing bourgeoisie. Advocates of this proposition would have to produce evidence to the effect that the observed rate of increase can be accounted for by a sequence of favorable circumstances unconnected with the mechanism of private enterprise and strong enough to overcome the latter's resistance. This is precisely the question which we shall discuss in Chapter IX. However, those who espouse this variant at least avoid the trouble about historical fact that the advocates of the alternative proposition have to face. This avers that capitalist reality once tended to favor maximum productive performance, or at all events productive performance so considerable as to constitute a major element in any serious appraisal of the system; but that the later spread of monopolist structures, killing competition, has by now reversed that tendency.

First, this involves the creation of an entirely imaginary golden age of perfect competition that at some time somehow metamorphosed itself into the monopolistic age, whereas it is quite clear that perfect competition has at no time been more of a reality than it is at present. Secondly, it is necessary to point out that the rate of increase in output did not decrease from the nineties from which, I suppose, the prevalence of the largest-size concerns, at least in manufacturing industry, would have to be dated; that there is nothing in the behavior of the time series of total output to suggest a "break in trend"; and, most important of all, that the modern standard of life of the masses evolved during the period of relatively unfettered "big business." If we list the items that enter the modern workman's budget and from 1899 on observe the course of their prices not in terms of money but in terms of the hours of labor that will buy them—i.e., each year's money prices divided by each year's hourly wage rates—we cannot fail to be struck by the rate of the advance which, considering the spectacular improvement in qualities, seems to have been greater and not smaller than it ever was before. If we economists were given less to wishful thinking and more to the observation of facts, doubts would

immediately arise as to the realistic virtues of a theory that would have led us to expect a very different result. Nor is this all. As soon as we go into details and inquire into the individual items in which progress was most conspicuous, the trail leads not to the doors of those firms that work under conditions of comparatively free competition but precisely to the doors of the large concerns—which, as in the case of agricultural machinery, also account for much of the progress in the competitive sector—and a shocking suspicion dawns upon us that big business may have had more to do with creating that standard of life than with keeping it down.

The conclusions alluded to at the end of the preceding chapter are in fact almost completely false. Yet they follow from observations and theorems that are almost completely<sup>1</sup> true. Both economists and popular writers have once more run away with some fragments of reality they happened to grasp. These fragments themselves were mostly seen correctly. Their formal properties were mostly developed correctly. But no conclusions about capitalist reality as a whole follow from such fragmentary analyses. If we draw them nevertheless, we can be right only by accident. That has been done. And the lucky accident did not happen.

The essential point to grasp is that in dealing with capitalism we are dealing with an evolutionary process. It may seem strange that anyone can fail to see so obvious a fact which moreover was long ago emphasized by Karl Marx. Yet that fragmentary analysis which yields the bulk of our propositions about the functioning of modern capitalism persistently neglects it. Let us restate the point and see how it bears upon our problem.

Capitalism, then, is by nature a form or method of economic change and not only never is but never can be stationary. And this evolutionary character of the capitalist process is not merely due to the fact that economic life goes on in a social and natural environment which changes and by its change alters the data of economic action; this fact is important and these changes (wars, revolutions and so on) often condition industrial change, but they are not its prime movers. Nor is this evolutionary character due to a quasi-automatic increase in population and capital or to the vagaries of monetary systems<sup>2</sup> of

<sup>1</sup> As a matter of fact, those observations and theorems are not completely satisfactory. The usual expositions of the doctrine of imperfect competition fail in particular to give due attention to the many and important cases in which, even as a matter of static theory, imperfect competition approximates the results of perfect competition. There are other cases in which it does not do this, but offers compensations which, while not entering any output index, yet contribute to what the output index is in the last resort intended to measure—the cases in which a firm defends its market by establishing a name for quality and service for instance. However, in order to simplify matters, we will not take issue with that doctrine on its own ground.

which exactly the same thing holds true. The fundamental impulse that sets and keeps the capitalist engine in motion comes from the new consumers' goods, the new methods of production or transportation, the new markets, the new forms of industrial organization that capitalist enterprise creates.

As we have seen in the preceding chapter, the contents of the laborer's budget, say from 1760 to 1940, did not simply grow on unchanging lines but they underwent a process of qualitative change. Similarly, the history of the productive apparatus of a typical farm, from the beginnings of the rationalization of crop rotation, plowing and fattening to the mechanized thing of today—linking up with elevators and railroads—is a history of revolutions. So is the history of the productive apparatus of the iron and steel industry from the charcoal furnace to our own type of furnace, or the history of the apparatus of power production from the overshot water wheel to the modern power plant, or the history of transportation from the mail-coach to the airplane. The opening up of new markets, foreign or domestic, and the organizational development from the craft shop and factory to such concerns as U. S. Steel illustrate the same process of industrial mutation—if I may use that biological term—that incessantly revolutionizes<sup>2</sup> the economic structure *from within*, incessantly destroying the old one, incessantly creating a new one. This process of Creative Destruction is the essential fact about capitalism. It is what capitalism consists in and what every capitalist concern has got to live in. This fact bears upon our problem in two ways.

First, since we are dealing with a process whose every element takes considerable time in revealing its true features and ultimate effects, there is no point in appraising the performance of that process *ex visu* of a given point of time; we must judge its performance over time, as it unfolds through decades or centuries. A system—any system, economic or other—that at every given point of time fully utilizes its possibilities to the best advantage may yet in the long run be inferior to a system that does so at no given point of time, because the latter's failure to do so may be a condition for the level or speed of long-run performance.

Second, since we are dealing with an organic process, analysis of what happens in any particular part of it—say, in an individual concern or industry—may indeed clarify details of mechanism but is inconclusive beyond that. Every piece of business strategy acquires its true significance only against the background of that process and

<sup>2</sup> Those revolutions are not strictly incessant; they occur in discrete rushes which are separated from each other by spans of comparative quiet. The process as a whole works incessantly however, in the sense that there always is either revolution or absorption of the results of revolution, both together forming what are known as business cycles.

within the situation created by it. It must be seen in its role in the perennial gale of creative destruction; it cannot be understood irrespective of it or, in fact, on the hypothesis that there is a perennial lull.

But economists who, *ex visu* of a point of time, look for example at the behavior of an oligopolist industry—an industry which consists of a few big firms—and observe the well-known moves and countermoves within it that seem to aim at nothing but high prices and restrictions of output are making precisely that hypothesis. They accept the data of the momentary situation as if there were no past or future to it and think that they have understood what there is to understand if they interpret the behavior of those firms by means of the principle of maximizing profits with reference to those data. The usual theorist's paper and the usual government commission's report practically never try to see that behavior, on the one hand, as a result of a piece of past history and, on the other hand, as an attempt to deal with a situation that is sure to change presently—as an attempt by those firms to keep on their feet, on ground that is slipping away from under them. In other words, the problem that is usually being visualized is how capitalism administers existing structures, whereas the relevant problem is how it creates and destroys them. As long as this is not recognized, the investigator does a meaningless job. As soon as it is recognized, his outlook on capitalist practice and its social results changes considerably.<sup>3</sup>

The first thing to go is the traditional conception of the *modus operandi* of competition. Economists are at long last emerging from the stage in which price competition was all they saw. As soon as quality competition and sales effort are admitted into the sacred precincts of theory, the price variable is ousted from its dominant position. However, it is still competition within a rigid pattern of invariant conditions, methods of production and forms of industrial organization in particular, that practically monopolizes attention. But in capitalist reality as distinguished from its textbook picture, it is not that kind of competition which counts but the competition from the new commodity, the new technology, the new source of supply, the new type of organization (the largest-scale unit of control for instance)—competition which commands a decisive cost or quality advantage and which strikes not at the margins of the profits and the outputs of the existing firms but at their foundations and their very lives. This kind of competition is as much more effective than the other as a bombardment is in comparison with forcing a door, and

<sup>3</sup> It should be understood that it is only our appraisal of economic performance and not our moral judgment that can be so changed. Owing to its autonomy, moral approval or disapproval is entirely independent of our appraisal of social (or any other) results, unless we happen to adopt a moral system such as utilitarianism

which makes moral approval and disapproval turn on them *ex definitione*.

so much more important that it becomes a matter of comparative indifference whether competition in the ordinary sense functions more or less promptly; the powerful lever that in the long run expands output and brings down prices is in any case made of other stuff.

It is hardly necessary to point out that competition of the kind we now have in mind acts not only when in being but also when it is merely an ever-present threat. It disciplines before it attacks. The businessman feels himself to be in a competitive situation even if he is alone in his field or if, though not alone, he holds a position such that investigating government experts fail to see any effective competition between him and any other firms in the same or a neighboring field and in consequence conclude that his talk, under examination, about his competitive sorrows is all make-believe. In many cases, though not in all, this will in the long run enforce behavior very similar to the perfectly competitive pattern.

Many theorists take the opposite view which is best conveyed by an example. Let us assume that there is a certain number of retailers in a neighborhood who try to improve their relative position by service and "atmosphere" but avoid price competition and stick as to methods to the local tradition—a picture of stagnating routine. As others drift into the trade that quasi-equilibrium is indeed upset, but in a manner that does not benefit their customers. The economic space around each of the shops having been narrowed, their owners will no longer be able to make a living and they will try to mend the case by raising prices in tacit agreement. This will further reduce their sales and so, by successive pyramiding, a situation will evolve in which increasing potential supply will be attended by increasing instead of decreasing prices and by decreasing instead of increasing sales.

Such cases do occur, and it is right and proper to work them out. But as the practical instances usually given show, they are fringe-end cases to be found mainly in the sectors furthest removed from all that is most characteristic of capitalist activity.<sup>4</sup> Moreover, they are transient by nature. In the case of retail trade the competition that matters arises not from additional shops of the same type, but from the department store, the chain store, the mail-order house and the super-market which are bound to destroy those pyramids sooner or later.<sup>5</sup>

<sup>4</sup> This is also shown by a theorem we frequently meet with in expositions of the theory of imperfect competition, viz., the theorem that, under conditions of imperfect competition, producing or trading businesses tend to be irrationally small. Since imperfect competition is at the same time held to be an outstanding characteristic of modern industry we are set to wondering what world these theorists live in, unless, as stated above, fringe-end cases are all they have in mind.

<sup>5</sup> The mere threat of their attack cannot, in the particular conditions, environmental and personal, of small-scale retail trade, have its usual disciplining influence, for the small man is too much hampered by his cost structure and, however

Now a theoretical construction which neglects this essential element of the case neglects all that is most typically capitalist about it; even if correct in logic as well as in fact, it is like *Hamlet* without the Danish prince.

well he may manage within his inescapable limitations, he can never adapt himself to the methods of competitors who can afford to sell at the price at which he buys.

## CHAPTER VIII

## MONOPOLISTIC PRACTICES

WHAT has been said so far is really sufficient to enable the reader to deal with the large majority of the practical cases he is likely to meet and to realize the inadequacy of most of those criticisms of the profit economy which, directly or indirectly, rely on the absence of perfect competition. Since, however, the bearing of our argument on some of those criticisms may not be obvious at a glance, it will be worth our while to elaborate a little in order to make a few points more explicit.

1. We have just seen that, both as a fact and as a threat, the impact of new things—new technologies for instance—on the existing structure of an industry considerably reduces the long-run scope and importance of practices that aim, through restricting output, at conserving established positions and at maximizing the profits accruing from them. We must now recognize the further fact that restrictive practices of this kind, as far as they are effective, acquire a new significance in the perennial gale of creative destruction, a significance which they would not have in a stationary state or in a state of slow and balanced growth. In either of these cases restrictive strategy would produce no result other than an increase in profits at the expense of buyers except that, in the case of balanced advance, it might still prove to be the easiest and most effective way of collecting the means by which to finance additional investment.<sup>1</sup> But in the process of creative destruction, restrictive practices may do much to steady the ship and to alleviate temporary difficulties. This is in fact a very familiar argument which always turns up in times of depression and, as everyone knows, has become very popular with governments and their economic advisers—witness the NRA. While it has been so much misused and so faultily acted upon that most economists heartily despise it, those

<sup>1</sup>Theorists are apt to look upon anyone who admits this possibility as guilty of gross error, and to prove immediately that financing by borrowing from banks or from private savers or, in the case of public enterprise, financing from the proceeds of an income tax is much more rational than is financing from surplus profits collected through a restrictive policy. For some patterns of behavior they are quite right. For others they are quite wrong. I believe that both capitalism and communism of the Russian type belong in the latter category. But the point is that theoretical considerations, especially theoretical considerations of the short-run kind, cannot solve, although they contribute to the solution of, the problem which we shall meet again in the next part.

# Institut Hayek

*Libéralisme & atlantisme*

27-9-2003

**The Use of Knowledge in Society**

F.A. Hayek

---

## **The Use of Knowledge in Society**

**F.A. Hayek**

*American Economic Review*, XXXV, No. 4; September, 1945, pp. 519-30.

### **I**

What is the problem we wish to solve when we try to construct a rational economic order? On certain familiar assumptions the answer is simple enough. *If* we possess all the relevant information, *if* we can start out from a given system of preferences, and *if* we command complete knowledge of available means, the problem which remains is purely one of logic. That is, the answer to the question of what is the best use of the available means is implicit in our assumptions. The conditions which the solution of this optimum problem must satisfy have been fully worked out and can be stated best in mathematical form: put at their briefest, they are that the marginal rates of substitution between any two commodities or factors must be the same in all their different uses. H.1

This, however, is emphatically *not* the economic problem which society faces. And the economic calculus which we have developed to solve this logical problem, though an important step toward the solution of the economic problem of society, does not yet provide an answer to it. The reason for this is that the "data" from which the economic calculus starts are never for the whole society "given" to a single mind which could work out the implications and can never be so given. H.2

The peculiar character of the problem of a rational economic order is determined precisely by the fact that the knowledge of the circumstances of which we must make use never exists in concentrated or integrated form but solely as the dispersed bits of incomplete and frequently contradictory knowledge which all the separate individuals possess. The economic problem of society is thus not merely a problem of how to allocate "given" resources—if "given" is taken to mean given to a single mind which deliberately solves the problem set by these "data." It is rather a problem of how to secure the best use of resources known to any of the members of society, for ends whose relative importance only these individuals know. Or, to put it briefly, it is a problem of the utilization of knowledge which is not given to anyone in its totality. H.3

This character of the fundamental problem has, I am afraid, been obscured rather than illuminated by many of the recent refinements of economic theory, particularly by many of the uses made of mathematics. Though the problem with which I want primarily to deal in this paper is the problem of a

rational economic organization, I shall in its course be led again and again to point to its close connections with certain methodological questions. Many of the points I wish to make are indeed conclusions toward which diverse paths of reasoning have unexpectedly converged. But, as I now see these problems, this is no accident. It seems to me that many of the current disputes with regard to both economic theory and economic policy have their common origin in a misconception about the nature of the economic problem of society. This misconception in turn is due to an erroneous transfer to social phenomena of the habits of thought we have developed in dealing with the phenomena of nature. H.4

## II

In ordinary language we describe by the word "planning" the complex of interrelated decisions about the allocation of our available resources. All economic activity is in this sense planning; and in any society in which many people collaborate, this planning, whoever does it, will in some measure have to be based on knowledge which, in the first instance, is not given to the planner but to somebody else, which somehow will have to be conveyed to the planner. The various ways in which the knowledge on which people base their plans is communicated to them is the crucial problem for any theory explaining the economic process, and the problem of what is the best way of utilizing knowledge initially dispersed among all the people is at least one of the main problems of economic policy—or of designing an efficient economic system. H.5

The answer to this question is closely connected with that other question which arises here, that of *who* is to do the planning. It is about this question that all the dispute about "economic planning" centers. This is not a dispute about whether planning is to be done or not. It is a dispute as to whether planning is to be done centrally, by one authority for the whole economic system, or is to be divided among many individuals. Planning in the specific sense in which the term is used in contemporary controversy necessarily means central planning—direction of the whole economic system according to one unified plan. Competition, on the other hand, means decentralized planning by many separate persons. The halfway house between the two, about which many people talk but which few like when they see it, is the delegation of planning to organized industries, or, in other words, monopoly. H.6

Which of these systems is likely to be more efficient depends mainly on the question under which of them we can expect that fuller use will be made of the existing knowledge. And this, in turn, depends on whether we are more likely to succeed in putting at the disposal of a single central authority all the knowledge which ought to be used but which is initially dispersed among many different individuals, or in conveying to the individuals such additional knowledge as they need in order to enable them to fit their plans with those of others. H.7

## III

It will at once be evident that on this point the position will be different with respect to different kinds of knowledge; and the answer to our question will therefore largely turn on the relative importance of the different kinds of knowledge; those more likely to be at the disposal of particular individuals and those which we should with greater confidence expect to find in the possession of an authority made up of suitably chosen experts. If it is today so widely assumed that the latter will be in a better position, this is because one kind of knowledge, namely, scientific knowledge, occupies now so prominent a place in public imagination that we tend to forget that it is not the only kind that is relevant. It may be admitted that, as far as scientific knowledge is concerned, a body of suitably chosen experts may be in the best position to command all the best knowledge available—though this is of course merely shifting the difficulty to the problem of selecting the experts. What I wish to point out is that, even assuming that this problem can be readily solved, it is only a small part of the wider problem. H.8

Today it is almost heresy to suggest that scientific knowledge is not the sum of all knowledge. But a little reflection will show that there is beyond question a body of very important but unorganized knowledge which cannot possibly be called scientific in the sense of knowledge of general rules: the knowledge of the particular circumstances of time and place. It is with respect to this that practically every individual has some advantage over all others because he possesses unique information of which beneficial use might be made, but of which use can be made only if the decisions depending on it are left to him or are made with his active coöperation. We need to remember only how much we have to learn in any occupation after we have completed our theoretical training, how big a part of our working life we spend learning particular jobs, and how valuable an asset in all walks of life is knowledge of people, of local conditions, and of special circumstances. To know of and put to use a machine not fully employed, or somebody's skill which could be better utilized, or to be aware of a surplus stock which can be drawn upon during an interruption of supplies, is socially quite as useful as the knowledge of better alternative techniques. And the shipper who earns his living from using otherwise empty or half-filled journeys of tramp-steamers, or the estate agent whose whole knowledge is almost exclusively one of temporary opportunities, or the *arbitrageur* who gains from local differences of commodity prices, are all performing eminently useful functions based on special knowledge of circumstances of the fleeting moment not known to others. H.9

It is a curious fact that this sort of knowledge should today be generally regarded with a kind of contempt and that anyone who by such knowledge gains an advantage over somebody better equipped with theoretical or technical knowledge is thought to have acted almost disreputably. To gain an advantage from better knowledge of facilities of communication or transport is sometimes regarded as almost dishonest, although it is quite as important that society make use of the best opportunities in this respect as in using the latest scientific discoveries. This prejudice has in a considerable measure affected the attitude toward commerce in general compared with that toward production. Even economists who regard themselves as definitely immune to the crude materialist fallacies of the past constantly commit the same mistake where activities directed toward the acquisition of such practical knowledge are concerned—apparently because in their scheme of things all such knowledge is supposed to be "given." The common idea now seems to be that all such knowledge should as a matter of course be readily at the command of everybody, and the reproach of irrationality leveled against the existing economic order is frequently based on the fact that it is not so available. This view disregards the fact that the method by which such knowledge can be made as widely available as possible is precisely the problem to which we have to find an answer. H.10

#### IV

If it is fashionable today to minimize the importance of the knowledge of the particular circumstances of time and place, this is closely connected with the smaller importance which is now attached to change as such. Indeed, there are few points on which the assumptions made (usually only implicitly) by the "planners" differ from those of their opponents as much as with regard to the significance and frequency of changes which will make substantial alterations of production plans necessary. Of course, if detailed economic plans could be laid down for fairly long periods in advance and then closely adhered to, so that no further economic decisions of importance would be required, the task of drawing up a comprehensive plan governing all economic activity would be much less formidable. H.11

It is, perhaps, worth stressing that economic problems arise always and only in consequence of change. So long as things continue as before, or at least as they were expected to, there arise no new problems requiring a decision, no need to form a new plan. The belief that changes, or at least day-to-day adjustments, have become less important in modern times implies the contention that economic problems

also have become less important. This belief in the decreasing importance of change is, for that reason, usually held by the same people who argue that the importance of economic considerations has been driven into the background by the growing importance of technological knowledge. H.12

Is it true that, with the elaborate apparatus of modern production, economic decisions are required only at long intervals, as when a new factory is to be erected or a new process to be introduced? Is it true that, once a plant has been built, the rest is all more or less mechanical, determined by the character of the plant, and leaving little to be changed in adapting to the ever-changing circumstances of the moment? H.13

The fairly widespread belief in the affirmative is not, as far as I can ascertain, borne out by the practical experience of the businessman. In a competitive industry at any rate—and such an industry alone can serve as a test—the task of keeping cost from rising requires constant struggle, absorbing a great part of the energy of the manager. How easy it is for an inefficient manager to dissipate the differentials on which profitability rests, and that it is possible, with the same technical facilities, to produce with a great variety of costs, are among the commonplaces of business experience which do not seem to be equally familiar in the study of the economist. The very strength of the desire, constantly voiced by producers and engineers, to be allowed to proceed untrammelled by considerations of money costs, is eloquent testimony to the extent to which these factors enter into their daily work. H.14

One reason why economists are increasingly apt to forget about the constant small changes which make up the whole economic picture is probably their growing preoccupation with statistical aggregates, which show a very much greater stability than the movements of the detail. The comparative stability of the aggregates cannot, however, be accounted for—as the statisticians occasionally seem to be inclined to do—by the "law of large numbers" or the mutual compensation of random changes. The number of elements with which we have to deal is not large enough for such accidental forces to produce stability. The continuous flow of goods and services is maintained by constant deliberate adjustments, by new dispositions made every day in the light of circumstances not known the day before, by *B* stepping in at once when *A* fails to deliver. Even the large and highly mechanized plant keeps going largely because of an environment upon which it can draw for all sorts of unexpected needs; tiles for its roof, stationery for its forms, and all the thousand and one kinds of equipment in which it cannot be self-contained and which the plans for the operation of the plant require to be readily available in the market. H.15

This is, perhaps, also the point where I should briefly mention the fact that the sort of knowledge with which I have been concerned is knowledge of the kind which by its nature cannot enter into statistics and therefore cannot be conveyed to any central authority in statistical form. The statistics which such a central authority would have to use would have to be arrived at precisely by abstracting from minor differences between the things, by lumping together, as resources of one kind, items which differ as regards location, quality, and other particulars, in a way which may be very significant for the specific decision. It follows from this that central planning based on statistical information by its nature cannot take direct account of these circumstances of time and place and that the central planner will have to find some way or other in which the decisions depending on them can be left to the "man on the spot." H.16

## V

If we can agree that the economic problem of society is mainly one of rapid adaptation to changes in the particular circumstances of time and place, it would seem to follow that the ultimate decisions must be left to the people who are familiar with these circumstances, who know directly of the relevant changes and of the resources immediately available to meet them. We cannot expect that this problem will be solved by first communicating all this knowledge to a central board which, after integrating *all*

knowledge, issues its orders. We must solve it by some form of decentralization. But this answers only part of our problem. We need decentralization because only thus can we insure that the knowledge of the particular circumstances of time and place will be promptly used. But the "man on the spot" cannot decide solely on the basis of his limited but intimate knowledge of the facts of his immediate surroundings. There still remains the problem of communicating to him such further information as he needs to fit his decisions into the whole pattern of changes of the larger economic system. H.17

How much knowledge does he need to do so successfully? Which of the events which happen beyond the horizon of his immediate knowledge are of relevance to his immediate decision, and how much of them need he know? H.18

There is hardly anything that happens anywhere in the world that *might* not have an effect on the decision he ought to make. But he need not know of these events as such, nor of *all* their effects. It does not matter for him *why* at the particular moment more screws of one size than of another are wanted, *why* paper bags are more readily available than canvas bags, or *why* skilled labor, or particular machine tools, have for the moment become more difficult to obtain. All that is significant for him is *how much more or less* difficult to procure they have become compared with other things with which he is also concerned, or how much more or less urgently wanted are the alternative things he produces or uses. It is always a question of the relative importance of the particular things with which he is concerned, and the causes which alter their relative importance are of no interest to him beyond the effect on those concrete things of his own environment. H.19

It is in this connection that what I have called the "economic calculus" proper helps us, at least by analogy, to see how this problem can be solved, and in fact is being solved, by the price system. Even the single controlling mind, in possession of all the data for some small, self-contained economic system, would not—every time some small adjustment in the allocation of resources had to be made—go explicitly through all the relations between ends and means which might possibly be affected. It is indeed the great contribution of the pure logic of choice that it has demonstrated conclusively that even such a single mind could solve this kind of problem only by constructing and constantly using rates of equivalence (or "values," or "marginal rates of substitution"), *i.e.*, by attaching to each kind of scarce resource a numerical index which cannot be derived from any property possessed by that particular thing, but which reflects, or in which is condensed, its significance in view of the whole means-end structure. In any small change he will have to consider only these quantitative indices (or "values") in which all the relevant information is concentrated; and, by adjusting the quantities one by one, he can appropriately rearrange his dispositions without having to solve the whole puzzle *ab initio* or without needing at any stage to survey it at once in all its ramifications. H.20

Fundamentally, in a system in which the knowledge of the relevant facts is dispersed among many people, prices can act to coordinate the separate actions of different people in the same way as subjective values help the individual to coordinate the parts of his plan. It is worth contemplating for a moment a very simple and commonplace instance of the action of the price system to see what precisely it accomplishes. Assume that somewhere in the world a new opportunity for the use of some raw material, say, tin, has arisen, or that one of the sources of supply of tin has been eliminated. It does not matter for our purpose—and it is very significant that it does not matter—which of these two causes has made tin more scarce. All that the users of tin need to know is that some of the tin they used to consume is now more profitably employed elsewhere and that, in consequence, they must economize tin. There is no need for the great majority of them even to know where the more urgent need has arisen, or in favor of what other needs they ought to husband the supply. If only some of them know directly of the new demand, and switch resources over to it, and if the people who are aware of the new gap thus created in turn fill it from still other sources, the effect will rapidly spread throughout the whole economic system and influence not

only all the uses of tin but also those of its substitutes and the substitutes of these substitutes, the supply of all the things made of tin, and their substitutes, and so on; and all this without the great majority of those instrumental in bringing about these substitutions knowing anything at all about the original cause of these changes. The whole acts as one market, not because any of its members survey the whole field, but because their limited individual fields of vision sufficiently overlap so that through many intermediaries the relevant information is communicated to all. The mere fact that there is one price for any commodity—or rather that local prices are connected in a manner determined by the cost of transport, etc.—brings about the solution which (it is just conceptually possible) might have been arrived at by one single mind possessing all the information which is in fact dispersed among all the people involved in the process. H.21

## VI

We must look at the price system as such a mechanism for communicating information if we want to understand its real function—a function which, of course, it fulfils less perfectly as prices grow more rigid. (Even when quoted prices have become quite rigid, however, the forces which would operate through changes in price still operate to a considerable extent through changes in the other terms of the contract.) The most significant fact about this system is the economy of knowledge with which it operates, or how little the individual participants need to know in order to be able to take the right action. In abbreviated form, by a kind of symbol, only the most essential information is passed on and passed on only to those concerned. It is more than a metaphor to describe the price system as a kind of machinery for registering change, or a system of telecommunications which enables individual producers to watch merely the movement of a few pointers, as an engineer might watch the hands of a few dials, in order to adjust their activities to changes of which they may never know more than is reflected in the price movement. H.22

Of course, these adjustments are probably never "perfect" in the sense in which the economist conceives of them in his equilibrium analysis. But I fear that our theoretical habits of approaching the problem with the assumption of more or less perfect knowledge on the part of almost everyone has made us somewhat blind to the true function of the price mechanism and led us to apply rather misleading standards in judging its efficiency. The marvel is that in a case like that of a scarcity of one raw material, without an order being issued, without more than perhaps a handful of people knowing the cause, tens of thousands of people whose identity could not be ascertained by months of investigation, are made to use the material or its products more sparingly; *i.e.*, they move in the right direction. This is enough of a marvel even if, in a constantly changing world, not all will hit it off so perfectly that their profit rates will always be maintained at the same constant or "normal" level. H.23

I have deliberately used the word "marvel" to shock the reader out of the complacency with which we often take the working of this mechanism for granted. I am convinced that if it were the result of deliberate human design, and if the people guided by the price changes understood that their decisions have significance far beyond their immediate aim, this mechanism would have been acclaimed as one of the greatest triumphs of the human mind. Its misfortune is the double one that it is not the product of human design and that the people guided by it usually do not know why they are made to do what they do. But those who clamor for "conscious direction"—and who cannot believe that anything which has evolved without design (and even without our understanding it) should solve problems which we should not be able to solve consciously—should remember this: The problem is precisely how to extend the span of out utilization of resources beyond the span of the control of any one mind; and therefore, how to dispense with the need of conscious control, and how to provide inducements which will make the individuals do the desirable things without anyone having to tell them what to do. H.24

The problem which we meet here is by no means peculiar to economics but arises in connection with nearly all truly social phenomena, with language and with most of our cultural inheritance, and constitutes really the central theoretical problem of all social science. As Alfred Whitehead has said in another connection, "It is a profoundly erroneous truism, repeated by all copy-books and by eminent people when they are making speeches, that we should cultivate the habit of thinking what we are doing. The precise opposite is the case. Civilization advances by extending the number of important operations which we can perform without thinking about them." This is of profound significance in the social field. We make constant use of formulas, symbols, and rules whose meaning we do not understand and through the use of which we avail ourselves of the assistance of knowledge which individually we do not possess. We have developed these practices and institutions by building upon habits and institutions which have proved successful in their own sphere and which have in turn become the foundation of the civilization we have built up. H.25

The price system is just one of those formations which man has learned to use (though he is still very far from having learned to make the best use of it) after he had stumbled upon it without understanding it. Through it not only a division of labor but also a coordinated utilization of resources based on an equally divided knowledge has become possible. The people who like to deride any suggestion that this may be so usually distort the argument by insinuating that it asserts that by some miracle just that sort of system has spontaneously grown up which is best suited to modern civilization. It is the other way round: man has been able to develop that division of labor on which our civilization is based because he happened to stumble upon a method which made it possible. Had he not done so, he might still have developed some other, altogether different, type of civilization, something like the "state" of the termite ants, or some other altogether unimaginable type. All that we can say is that nobody has yet succeeded in designing an alternative system in which certain features of the existing one can be preserved which are dear even to those who most violently assail it—such as particularly the extent to which the individual can choose his pursuits and consequently freely use his own knowledge and skill. H.26

## VII

It is in many ways fortunate that the dispute about the indispensability of the price system for any rational calculation in a complex society is now no longer conducted entirely between camps holding different political views. The thesis that without the price system we could not preserve a society based on such extensive division of labor as ours was greeted with a howl of derision when it was first advanced by von Mises twenty-five years ago. Today the difficulties which some still find in accepting it are no longer mainly political, and this makes for an atmosphere much more conducive to reasonable discussion. When we find Leon Trotsky arguing that "economic accounting is unthinkable without market relations"; when Professor Oscar Lange promises Professor von Mises a statue in the marble halls of the future Central Planning Board; and when Professor Abba P. Lerner rediscovers Adam Smith and emphasizes that the essential utility of the price system consists in inducing the individual, while seeking his own interest, to do what is in the general interest, the differences can indeed no longer be ascribed to political prejudice. The remaining dissent seems clearly to be due to purely intellectual, and more particularly methodological, differences. H.27

A recent statement by Professor Joseph Schumpeter in his *Capitalism, Socialism, and Democracy* provides a clear illustration of one of the methodological differences which I have in mind. Its author is pre-eminent among those economists who approach economic phenomena in the light of a certain branch of positivism. To him these phenomena accordingly appear as objectively given quantities of commodities impinging directly upon each other, almost, it would seem, without any intervention of human minds. Only against this background can I account for the following (to me startling) pronouncement. Professor Schumpeter argues that the possibility of a rational calculation in the absence of markets for the factors of

production follows for the theorist "from the elementary proposition that consumers in evaluating ('demanding') consumers' goods *ipso facto* also evaluate the means of production which enter into the production of these goods."\*1 H.28

Taken literally, this statement is simply untrue. The consumers do nothing of the kind. What Professor Schumpeter's "*ipso facto*" presumably means is that the valuation of the factors of production is implied in, or follows necessarily from, the valuation of consumers' goods. But this, too, is not correct. Implication is a logical relationship which can be meaningfully asserted only of propositions simultaneously present to one and the same mind. It is evident, however, that the values of the factors of production do not depend solely on the valuation of the consumers' goods but also on the conditions of supply of the various factors of production. Only to a mind to which all these facts were simultaneously known would the answer necessarily follow from the facts given to it. The practical problem, however, arises precisely because these facts are never so given to a single mind, and because, in consequence, it is necessary that in the solution of the problem knowledge should be used that is dispersed among many people. H.29

The problem is thus in no way solved if we can show that all the facts, *if* they were known to a single mind (as we hypothetically assume them to be given to the observing economist), would uniquely determine the solution; instead we must show how a solution is produced by the interactions of people each of whom possesses only partial knowledge. To assume all the knowledge to be given to a single mind in the same manner in which we assume it to be given to us as the explaining economists is to assume the problem away and to disregard everything that is important and significant in the real world. H.30

That an economist of Professor Schumpeter's standing should thus have fallen into a trap which the ambiguity of the term "datum" sets to the unwary can hardly be explained as a simple error. It suggests rather that there is something fundamentally wrong with an approach which habitually disregards an essential part of the phenomena with which we have to deal: the unavoidable imperfection of man's knowledge and the consequent need for a process by which knowledge is constantly communicated and acquired. Any approach, such as that of much of mathematical economics with its simultaneous equations, which in effect starts from the assumption that people's *knowledge* corresponds with the objective *facts* of the situation, systematically leaves out what is our main task to explain. I am far from denying that in our system equilibrium analysis has a useful function to perform. But when it comes to the point where it misleads some of our leading thinkers into believing that the situation which it describes has direct relevance to the solution of practical problems, it is high time that we remember that it does not deal with the social process at all and that it is no more than a useful preliminary to the study of the main problem.

H.31

1. J. Schumpeter, *Capitalism, Socialism, and Democracy* (New York; Harper, 1942), p. 175. Professor Schumpeter is, I believe, also the original author of the myth that Pareto and Barone have "solved" the problem of socialist calculation. What they, and many others, did was merely to state the conditions which a rational allocation of resources would have to satisfy and to point out that these were essentially the same as the conditions of equilibrium of a competitive market. This is something altogether different from knowing how the allocation of resources satisfying these conditions can be found in practice. Pareto himself (from whom Barone has taken practically everything he has to say), far from claiming to have solved the practical problem, in fact explicitly denies that it can be solved without the help of the market. See his *Manuel d'économie pure* (2d ed., 1927), pp. 233-34. The relevant passage is quoted in an English translation at the beginning of my article on "Socialist Calculation: The Competitive 'Solution,'" in *Economica*, New Series, Vol. VIII, No. 26 (May, 1940), p. 125.

*American Economic Review*, XXXV, No. 4; September, 1945, pp. 519-30.

---

[Retour page d'accueil](#)

© Institut Hayek