Islamic Economic Studies
Vol. 3, No. 2, June 1996

RULES FOR BENEFICIAL PRIVATIZATION: PRACTICAL IMPLICATIONS OF ECONOMIC ANALYSIS

WILLIAM J. BAUMOL*

In market economies, privatization of larger enterprises has been followed by regulation to prevent the firm from exercising market power and to enable competitors to enter the market and survive. Often, the regulators' rules have been poorly designed, raising costs, indirectly impeding competition and removing incentives for innovation and productivity growth. Analysis of regulatory experience in the US and elsewhere enabled economists to devise rules that can serve the regulators' goals more effectively -- rules being adopted in a number of countries. This paper explains the nature of the regulatory issues and the associated problems. It also describes the economists' proposed regulatory rules, particularly the price cap *approach to prevention of monopoly profits while preserving incentives for efficiency and innovation, and the* parity pricing *rule for access to bottleneck facilities, that is, to facilities owned by a single firm without which no competitor can operate successfully.*

1. THE DANGERS OF INAPPROPRIATE REGULATORY RULES FOR PRIVATIZED ENTERPRISE

From a view widespread only a few years ago that nationalization was the one way to protect public interest when suppliers possess monopoly power or when their products are, in some sense, indispensable, the pendulum has swung to the opposite position. The common wisdom has now leaped to the conclusion that only private enterprise can be relied upon to innovate, to produce efficiently and to adapt itself with dedication to the desires of the public. With this pro-private-enterprise attitude has come a movement toward reversal of what had been done before -- firms that had been taken over by government were to be returned to the private sector as quickly as was possible, and on the best terms that could be arrived at without excessive delay. This process has been undertaken not only in the formerly Communist countries but also in the West, encompassing Europe, Australasia and Latin America. It is a process that is still under way.

Practice has shown that some of the earlier expectations of privatization were excessively optimistic. The process has been found in many cases to take longer than had been anticipated. Inflated predictions about what privatization can be ex-

* Director, C.V. Starr Center for Applied Economics, New York University, and professor emeritus and senior research economist, Princeton University. The author is deeply grateful to the Alfred P. Sloan Foundation and the C.V. Starr Center for their support of this work.

pected to accomplish were predictably deflated by subsequent reality. Moreover, it transpired that insufficient thought had been given to the appropriate terms and conditions under which the privatized firms should operate. Here, I will focus on the last of these issues, the one for which economic analysis can provide the most direct illumination.

The discussion here will show that, though privatization can indeed stimulate productivity growth and innovation and help to tailor product specifications to the wishes of consumers, it will be most effective in achieving those goals only if government avoids indefensible forms of interference with managerial decisions, but if it *does*, at the same time, and under certain appropriate circumstances, impose some minimal constraints upon the behavior of the privatized firm.

Here, the nature of these constraints will be described, and their purpose and proper implementation will be discussed in some detail. The discussion will be based both on the lessons of the early experience in the operation of the privatized firms and the pertinent implications of pure economic analysis.

The fact is that privatization has proved to be less easy to carry out, and less magical in its accomplishments, than seems earlier to have been believed. Disappointment has been engendered by the slow pace of privatization, the reduction in jobs that accompanies it, the wealth accumulated by those who are successful at the task, and the apparently long delays before its benefits emerge. These difficulties, of course, are the result of illusion and misunderstanding by the public. They do not recognize that even in the miracle economies of the Far East, decades were required before the invested effort began to provide benefits of noticeable magnitude. Nor is it recognized that efficiency and competitiveness require elimination of redundant jobs. Perhaps most important, it is not generally understood that the market's incentive mechanism is founded upon the prospects of wealth for those who succeed. Never having recognized these less popular attributes of the market mechanism, the public is distressed when they occur.

Perhaps most disturbing to the observer of the process of privatization, however, is the propensity of even well-intentioned public servants to sabotage that process. It is they who are likely to have made the privatized firm into a monopoly. It is they who, believing that they are encouraging competition, create what are, in effect, governmentally sponsored groups of firms, in which there coexist many enterprises, each of which is prohibited from competing with the others, and in which the most inefficient of the firms in the industry are kept alive at the consumer’s expense by impediments to price reductions by more efficient rivals. It is the bureaucrats, who, paying lip service to the market mechanism, but distrusting it profoundly, seek to take away the power of the privatized firms to make decisions for themselves, under the constraining influence of market forces. Finally, it is they who are prone to restrict the profits of the privatized firms and the incomes of their entrepreneurs, without recognizing that they are thereby destroying the very engine that can in time yield the benefits so widely expected to flow from privatization.

1.1 Why Constrain Privatized Firms - And Why Not?

If privatization is carried out because the market is believed to be a more efficient guide to the decisions and behavior of the enterprise than governmental administrators can be, one may well ask why government interference with the activities of the firm after its transfer to the private sector should even be contemplated. A little consideration of the issue will readily suggest the answer. Typically, the firms that are transferred from government ownership are at first monopolies or, even if they are not, there is often reason to suspect that they possess market power, meaning that they have the ability to raise the prices of their products above competitive levels, and to keep them high for a considerable period without fear of loss of substantial amounts of business to competitors. This is not a mere accident of history. The privatized firms possess their market power either because of the criterion on which they had previously been selected for nationalization, or, more simply, because governmental decisions conferred that power upon them. We will return to this issue presently. It is appropriate, however, to precede this by noting some of the consequences.

The obvious objection to unconstrained operation of a privatized monopoly entails the usual list of damage to the public interest threatened by monopoly. It can lead to misallocation of resources as the enterprise restricts its output below the socially-optimal level in order to force the market to increase the price of its product. It consequently is likely to entail overcharging -- a direct burden upon consumers. The firm is also to be expected to take steps to preserve its monopoly power, using whatever means it can to prevent or discourage the entry of rivals. This can entail steps ranging from costly legal challenges that use the courts to make entry difficult, to unwillingness, to provide access to prospective entrants to facilities (often referred to as ‘bottleneck facilities’) that the monopolist alone possesses and without which the entrant cannot hope to succeed. The role of such bottlenecks is apt to prove of critical importance, and has been an issue that has recently received a good deal of attention in the courts and regulatory agencies of the US, Great Britain and elsewhere. I will consequently return to the issue in some detail in the next section.

The valid reason for regulation of the privatized firms is, then, obvious. It is to constrain that firm from exercising monopoly power in ways that are detrimental to the public interest - to keep it from doing any of the things that have just been listed. On the other hand, the design of rules to achieve this objective is an extraordinarily delicate matter. The rules must not be so weak as to fail in their purpose, permitting the monopoly to act almost as it would have in their absence. On the other hand, they must not be so strict as to undermine the purpose of the act of privatization. They must permit prices and the allocation of resources to proceed as they would in an unregulated competitive market, thus either allowing the market forces to govern the pertinent acts and decisions or forcing the firm to act as though

it were governed by competitive market forces, even though such forces are not really present or are insufficiently strong for the purpose. Yet the rules must also preserve considerable freedom of action for management, for, otherwise the firm will continue to be governed by the bureaucracy, and lose all of the pressures for economic efficiency that genuine freedom of enterprise can provide. Finally, while preventing the earning of monopoly profits, they must not set artificial limits on profits or restrict profits on the basis of mechanistic formulas that eliminate the incentive for efficiency, productivity growth and innovation. This, too, is a difficult requirement partly because no one knows how to determine what portion of a firm's profits is ascribable to monopoly power and what portion is attributable to outstanding performance in terms of productivity and product quality and improvement. Section 2 will, consequently, show how this issue can be approached in a manner that is promising and that is increasingly being adopted in industrialized economies.

Next, let me offer a few observations about the market mechanism -- about what it can accomplish and about the means the market uses as its tools in these achievements. These remarks are not intended to be yet another restatement of the economist's views on the virtues of the market. Rather, they will emphasize only matters that should be kept in mind in dealing with privatization.

1.2 The Market: Some of its Virtues and Mechanisms

Historical evidence offers strong support for those who view the market primarily as an engine of growth, placing less emphasis than economists do on the market's ability to achieve a static allocation of resources that is efficient. The estimates, crude as they necessarily are, suggest that, on the average, productivity growth between third-century Rome and eighteenth-century Britain was approximately zero. Even during the period of medieval and Renaissance growth since the tenth century, productivity must have progressed at a minuscule pace, for otherwise humanity, which obviously produced at least enough to permit bare survival at the beginning of the period, would, in the intervening seven centuries, have been able to get well beyond the indescribable poverty (punctuated by widespread and frequent years of death by starvation) that characterized a period even as late as the seventeenth century. Yet, in the 200 years that have followed the inception of the Industrial Revolution, labor productivity in the wealthier countries has risen somewhere in the neighborhood of twentyfold, and with something like a halving of number of hours of work per year, have raised real per capita incomes by perhaps a factor of ten.

How has the market mechanism produced these miracles? The answer is, of course, complex and far from certain. Yet, there are several elements that are particularly pertinent for privatization policy. The first of these is the fact that the market serves the consumer effectively and is able to produce its abundance by virtue of its merciless system of rewards and penalties. The management that fails to

perform to the standards set by competition, that incurs costs higher than those of its rivals or turns out products less attractive to consumers, or lags innovation, can expect only one fate -- bankruptcy and demise. The impersonal forces of the market are harsh and unforgiving. The market simply accepts no excuse for failure, and summary execution is its prescribed penalty.

The market is also discomfortingly generous in its rewards to those who pass its tests. It showers wealth upon those who succeed in producing what the public wants at a cost as low as is currently attainable. Here it makes no distinction in terms of merit. The firm that stumbles on a popular new product is rewarded as handsomely as the one that invests long and heavily in the product development process. Efficiency is repaid whether it is the result of superior ability, greater effort or happenstance. For, the public benefits when products are cheap and meet its desires, regardless of whether the low cost or superior design is the result of meritorious effort, accident or simple greed.

Greed is, as a matter of fact, the key to the efficiency of the market mechanism or, rather, it is greed constrained and channeled by competition. This is the miracle of Adam Smith's invisible hand, that has not only developed the means to prevent the greedy from exploiting others, but has actually harnessed their greed to serve the interests of the general public. Competition *forces* those in pursuit of income and wealth to do so by building better mousetraps and by doing so at lower cost. As we are told in the invisible hand passage,

> "[The individual], generally indeed, neither intends to promote the public interest nor knows how much he is promoting it....by directing...industry in such a manner as its produce may be of the greatest value, he intends only his own gain, and he is in this, as in many other cases, led by an invisible hand to promote an end which was no part of his intention. By pursuing his own interest he frequently promotes that of society more effectually than when he really intends to promote it. I have never known much good done by those who affected to trade for the public good. It is an affectation, indeed, not very common among merchants, and very few words need be employed in dissuading them from it." (*Wealth of Nations*, Book 4, Chapter 2).

These are attributes of the market that are not widely understood. The misapprehension tempts the public to clamor for the imposition of limits on the wealth of those who are most successful. They lead to political support for public sector rescue of failing enterprises and dying industries. Above all, they engender public contempt and even hatred for those who are patently driven by greed to conquer new economic worlds and expand the economic horizon. One cannot quarrel with such preferences, resting as they do on deep convictions. But one can deplore the public's failure to recognize that it cannot have things both ways; that enforcement of its moral precepts condemns the invisible hand to impotence, and all but de-

stroys the ability of the market mechanism to produce high and expanding living standards, prosperity and abundance. All too often, failure to understand these matters has led political leaders to pander to the indignation of the general public by the adoption of regulation that effectively precluded the earning of profits or subsequently expropriated those profits. Where this has happened or has even been threatened, disappointment in the performance of the market mechanism has followed rapidly. Elimination of the rewards of enterprise is the sure way to prevent such enterprise from being undertaken. Entrepreneurs, seeing the productive avenues to wealth, prestige and power closed off, can be relied upon to find other ways to pursue their goals, but those other means will often provide very limited benefits to the public, sometimes none at all, and sometimes they may succeed only at heavy cost to the public.[1]

1.3 The Tendency Towards Monopoly in Western Firms Picked for Privatizing

The erection of impediments to the productive contribution of the privatized enterprises is hardly fortuitous; on the contrary, there is reason to expect the problems to recur and persist. We have just seen how failure by politicians and the general public to understand the nature of the market mechanism can lead to such difficulties. But we noted earlier that there is another key source of difficulty. A considerable proportion of the candidates for privatization have been firms with monopoly attributes. There are at least four reasons for this phenomenon, none of them accidental.

(i) Monopolies as targets of previous nationalization

In the free-enterprise economies, monopolies and near monopolies have been prime targets of takeover by government, so that when it is proposed to return nationalized firms to the private sector it is not surprising that a considerable proportion of the group of candidates hold either a monopoly position or a position that confers some degree of monopoly power upon the enterprise. When the venerable market economies have nationalized firms deliberately (that is, when they did not simply become residual proprietors of failing private firms) they tended to do so when there seemed to be reason to doubt the ability of market forces to channel the activities of those firms in the directions called for by the public interest. Consequently, in these economies the nationalized firms are generally large and there is

[1] The explosion of criminal activity in the former Soviet Union--as evidenced by the growth of its own 'Mafia'-- is surely a predictable consequence of continuing interference with legitimate business activity and the suffocating bureaucratic rules and practices that make investment unattractive and success so difficult for anyone that wishes to invest in factories, mining or distribution facilities and use them to contribute to their nations' outputs.

often reason to suspect them of possessing market power. Public utilities are strongly represented in the group, for example.

(ii) Management's desire for protection from competition

A second powerful influence that skewed public enterprises toward monopoly has been the predictable behavior of their public sector managements. Most persons enthusiastically favor competition when it affects others, but not when it constrains themselves and invariably makes their lives more difficult. Bureaucrats are no exception, but they are in a better position than most to make the wish come true. A governmentally owned utility is often protected by law from all but the most insignificant forms of rivalry. The state telephone company, the post office and the electricity supplier are generally secured from the threat of entry, by government decree.

(iii) A monopoly's assets command higher prices

When a nationalized firm is put up for sale, those who are responsible for overseeing the transaction are likely to consider themselves obligated for the sake of the public interest to seek to obtain for the property as high a price as can be gotten. But it is obvious that higher price bids can be elicited if the property is offered along with a monopoly license that is protected against the entry of rivals. This temptation is sometimes too great to resist, with the consequence that the newly private firm enters the economy with the grant of a monopoly, automatically accompanied by the threat of close regulation in the future.

(iv) Competition undermines popular cross subsidies

A fourth and more subtle influence skews the nationalized firms toward possession of monopoly power. The agencies that run them, like the regulators in the US, have strongly and persistently favored a policy known in the business as 'universal service'. They are moved in this direction by an amalgam of natural inclination and political pressure. Particularly where the firm in question is a public utility and its product is widely desired, it has been considered bad policy to require classes of consumers whom it is particularly difficult to supply to pay the high cost of serving them. It was considered appropriate for the same price per unit to be offered to everyone -- to isolated farmers who could be reached by telephone only by long and underutilized stretches of wire, to inner-city firms to which mail delivery costs are increased by congestion, crime and other impediments, and to any other user whose service is extraordinarily difficult and expensive. More than that. Some nationalized services such as telecommunications are less costly to deliver to large firms, because of the scale economies available in such transactions, than they are to small-volume residential customers. Dedication to universal service impelled those

who control nationalized enterprises to price the residential services at levels considerably lower than the relative cost of serving those customers appeared to require. Characteristically, all this ended up in a complex system of cross subsidies, with the supplier of telephone service, for example, required to charge big business customers prices sufficiently high to yield the revenues that could cover losses incurred in serving isolated farmers or inner-city user firms or household customers in general.

Nothing in the preceding paragraph is to be construed to imply that these cross subsidies are indefensible or that their social purposes are unworthy. Rather, the point is that such cross subsidies are incompatible with competition and freedom of entry. An electricity supplier that overprices relative to the pertinent costs the power it supplies to large business customers, and supplies electricity to farmers at a loss, can continue to do so as long as it possesses an unchallenged monopoly. But if the field is opened to entrants those new rivals are likely to spring up soon enough, and to focus their efforts upon sales to the highly profitable business customers. The monopoly incumbent can be expected to denounce this selective entry strategy as 'cream skimming', but it is precisely what economists usually hope the entrants will do. The result, of course, is that the high profits of the business segment of the market will soon be forced down by competitive pressure, and the original incumbent will find itself deprived of the source of funds out of which it previously made up for the losses incurred in serving rural firms and households. The cross subsidy must come to an end either by choice of the original incumbent firm or as the result of its bankruptcy. To avoid this, those who set the rules for the operation of nationalized firms, like the regulators in the US, found a variety of reasons for the prohibition of competition and entry. This influence, too, contributed to the high frequency with which candidates for privatization turned out to be monopolies. It is an influence that, as we will see, continues to haunt the process of privatization, and even its sequel in the West.

1.4 The Tendency to Tight Regulation of the Privatized Firms

The result is that when a government enterprise is transferred to private ownership it often finds itself suspect. Its goals are often taken to be exploitation of the public and subversion of competition, and it is widely judged to have the power to attain those goals. The forces of the market are deemed, sometimes with good reason, to be too weak to constrain that enterprise adequately. Hence, private it may be permitted to become, but only under the heavy hand of regulation. Individuals are allowed to *own* it but they are given little opportunity to *control* it.

Regulation, as is well known, is hardly costless. Aside from the costly record keeping and paper work and other direct administrative burdens it engenders, it is likely to give rise to continuing and expensive litigation that diverts the attention of management; it is certain to delay the adoption and execution of decisions, so that the actions of the firm are condemned to lag behind changes in market conditions;

perhaps worst of all, incentives are likely to be distorted, so that the motivation for efficiency, innovation and proper pricing is attenuated if not altogether removed. Yet, in the circumstances, some regulation may prove unavoidable, even if the sale of the property has been carried out in an exemplary manner.

In practice, in Western economies, problems at this stage have arisen primarily from a different source -- inexperience with economic regulation, its pitfalls, and the practices that will keep its social costs within reasonable bounds. For the nations that have the nationalized firms to be privatized are obviously those which in the past have chosen nationalization over regulation as the instrument for control of monopoly power. Thus, it is hardly accidental that the privatizing economies are the ones least prepared by experience to institute and carry out a rational regulatory regime. In the process they have tended to want to learn for themselves, many of them possessing a very competent civil service, experienced and effective business managers and a group of highly qualified economists. Yet in a number of cases, they have simply repeated many of the mistakes of US regulation that it has taken decades to begin to ameliorate.

Perhaps the central error besetting the process has been what amounts to complete distrust of the market on the part of the novice regulators, even those who consider themselves to be avid partisans of the free enterprise system. They do believe that *elsewhere in the economy* the market does a good job of directing business activity in accord with the general welfare, but they seem to feel that the market loses all of those salutary powers to circumscribe the firm under the regulator's oversight. The regulatory agency resists attempts to offer any significant range of discretion to the management of the privatized firm in making its economic decisions.

Prices, accounting methods, perhaps investment and other decisions are constrained closely, so that the firm may find itself left with less freedom to act and the market with even less influence over those acts than they possessed when the enterprise was a property of the government. For then the firms were run by bureaucrats whose actions were supposedly driven by devotion to the public interest, while once privatized they are in the hands of individuals believed to be driven only by greed. Thus, the private owners, on this view, must be circumscribed even more narrowly than their public-sector predecessors had been. In these circumstances it should cause little surprise that the market provides few of the benefits to that industry that are so widely expected of it. For the market forces have for all practical purposes been exiled from the arena.

1.5 Recipes for Misguided Regulation

One encounters in some form in the regulation of the newly privatized enterprises virtually all of what economists consider to be the mistakes that long plagued regulation in the United States. Thus, the following list of questionable ac-

tions will seem familiar to those who have studied American regulatory history. These include (i) prevention or limitation of effective competition, (ii) ossification of cross subsidy, (iii) use of inefficient restrictions to protect competitors at the expense of competition, (iv) injection of costly and avoidable regulatory risk, (v) restriction of freedom of decision-making by management even within limits competitive conditions would permit, (vi) use of discredited criteria such as fully distributed cost for regulation of prices. This list is quite incomplete, but it is already sufficiently long to illustrate the point.

(i) Prevention or limitation of effective competition

The desire to nurture competition in the privatized arenas is, of course, the natural consequence of the fear of monopoly power that underlies the decision to regulate. Competition, if it can be introduced and expanded, is the natural way to put an end to monopoly power and to limit its exercise. But two problems beset this approach. First, the mere introduction of additional firms into the market is no guarantee of effective competition or of any competition at all, either if market or technological conditions such as scale economies impede or preclude it, or if regulatory restrictions all but prevent competition. Second, as we have just seen, one of the central problems that has plagued the adoption of rational regulatory policy has been the conflict between two of its goals -- the encouragement of competition versus the promotion of 'universal service' (i.e., the use of the products of the regulated firm by virtually all members of the consuming public). For effective competition is incompatible with retention of the cross subsidies[2] that are valued so highly by many regulators as the means to make universal service possible. This has been known to impel regulators to adopt rules that protect the cross subsidies by undermining or prohibiting competition.

Thus, in a recent report on privatization in the United Kingdom,[3] *The Economist* tells us that,

> "More subtly, the government has modified its policy. The original plan was to open [rail] passenger services to competition...[with] trains...run by franchised providers, offering competing services on each line....The government is backing away from that. Its fear is that the entrepreneurs would pick the best peak-time services. Off peak services would be left to British Rail, or disappear altogether.

[2] The term '*cross subsidy*' refers to the prices of the different commodities of a multiproduct firm. It can be defined, roughly, as a case where the price of one of the firm's products, call it product X, is too low to reimburse the firm for the costs that supply of X causes the firm to incur, but where the price of another product, Y, of that firm is sufficiently high to yield a surplus that covers the loss incurred by X.

[3] For a fuller and very illuminating discussion of the privatization process in the UK, see Johnson [1991, Chapter 5].

> So ....Only on a few routes will 'open access' (that is, competition) be allowed. Even in those cases, the core service will be provided by one operator who will be eligible for government subsidy. Any other operator running a train on the route will have to compete without subsidy....
>
> [Another] big privatization, that of the Post Office, could soon get a green light.... [But ministers] worry that privatization threatened the universal postal rate, which ensures that it costs the same to post a letter to any part of Britain." *(The Economist* [1993, p. 53]).

(ii) Ossification of cross subsidy

British Rail and the British Post Office are by no means the only organizations for which policy makers have undertaken to preserve the historic cross subsidies. In postal service the uniformity of charges, regardless of distance or cost of delivery, is widely considered sacred and one can hardly imagine a privatized post office anywhere that stands a good chance of being freed from this restriction. However, it applies to other arenas as well, and in some of these different approaches have been taken by the regulators. That is, they have not all sought to protect the monopoly or the monopoly power of the private firm, as is at least contemplated for some industries in the UK.

Thus, in New Zealand, when the telephone company, New Zealand Telecom, was transferred to private hands, a condition of the sale was that the firm taking over the company from the government continue the price advantages the nationalized predecessor firm had offered to residential subscribers. This stipulation, referred to as the 'Kiwi Share', is believed to entail losses in the supply of at least some of the residential services. That is, those services are said to bring in revenues at the Kiwi Share prices that fall short of the incremental costs of the services in question. Profitability of the enterprise then requires a cross subsidy from other customers of New Zealand Telecom, presumably the business subscriber. This has led to litigation with Clear Communications, the new rival of Telecom, as to whether Clear should somehow bear part of the cost burden. More to the point for the current discussion, such enshrined cross subsidies seem to have had marked effects on the prices of services other than those that the universal service goal seeks to promote, and those prices may well have been driven far out of line with those that economic efficiency requires.[4]

[4] In the US, adherence to the goal of 'universal service', with its accompanying cross subsidies, has eroded as deregulation has spread. It became clear in telecommunications, for example, that as entry occurred the cross subsidies would become unsustainable. Yet even here, as in a number of public utility arenas, some vestigial cross subsidy was retained. The suppliers, ostensibly voluntarily, agreed to supply what are called 'lifeline services', which offer the elderly or the

(iii) Imposed exclusive territories

Despite the fact that continued monopoly permits retention of the politically popular cross subsidies, the monopolistic character of many of the privatized firms has elicited a schizophrenic reaction from regulators. Since monopoly is accepted as an evil, many of them have undertaken to destroy it by the introduction of competitors into the regulated industry. But apparently driven by a desire, conscious or unconscious, to have it both ways, they have in at least some cases ended up with an arrangement entailing a multiplicity of firms as well as continued cross subsidy. In what *appears* to be a compromise they have carried out what in the US courts has been described as 'protecting competitors while undermining competition'. This they have done by imposing an arrangement upon the industry, one in which continued coexistence of two or more firms is ensured, but none is given the freedom to compete with the others in prices and related matters. Where this is done, cross subsidy need not be threatened, since each of the firms in the market can confidently offer a number of services at prices that exceed costs substantially and then provide the profits that can finance the desired subsidies to other services.

This is sometimes accomplished in subtle ways. For example, the price ceilings imposed on British Telecom have resulted in very low prices on rental of telephone lines for which the company felt forced to make up by means of high prices on number and duration of calls. Large business customers normally keep their lines very busy with calls, resulting in a high call/line ratio, so that this pricing arrangement made it difficult for British Telecom to compete for business customers with its relatively unregulated rival, Mercury. The call/line ratio pattern is reversed for residential and small business users, so that Telecom found itself with a considerable price advantage in this segment of the market. The result was virtually a split market, with Mercury in effect assigned the large-volume business customers, with near immunity from competition, while Telecom found itself in the same position in the residential market. It was as though Mercury had been assigned an exclusive license for operation in Scotland, and Telecom had received the same for Wales.

The net cost to society of imposition of a consortium arrangement is likely to be high. Such an exclusive-territories arrangement clearly does little or nothing to curtail monopoly power. In addition, it creates inefficiencies that a monopoly is likely to avoid. For any particular segment of the market may not happen to be assigned to the firm that can serve it at lowest cost. Moreover, there may be costly replica-

---

impoverished, or the residents of slum areas some basic services, with all luxury enhancements eliminated, at highly reduced prices. Because the magnitude of the cross subsidy is kept to moderate levels by this approach, and because several, if not all, of the suppliers of the services in question have more or less voluntarily followed it, it does not appear to have led regulators to try to restrict entry, and it apparently has not greatly affected the prices of other services. Still, political pressures have not permitted an end to regulatory intervention to preserve popular cross subsidies.

tion of facilities, and facilities that are withdrawn from service by each firm because of the limited market segment assigned to it may not be those that are the most inefficient in the industry. That is, plant A of firm X that is shut down may be more efficient than firm Y's plant B that continues in operation. Yet the regulator whose actions have created such an exclusive-territories arrangement is likely to congratulate himself for having injected competition into the arena without endangering universal service.

(iv) Imposition of avoidable regulatory risk

Risk is costly to firms and that cost is usually passed on to consumers, at least in part. The most obvious case of this is when the firm takes out insurance against the risk and the premiums are passed on to consumers in whole or in part. In addition to the risks that normally face an enterprise, the regulated firm faces the danger that regulators will change their minds and behave in a manner not foreseen by management. This means that managements may have made costly errors in their decisions, errors that could have been avoided if the regulators had given earlier notice of their future courses of action. This is true of all regulation but it affects privatized western firms in a distinctive way.

When the regulatory agency oversees the introduction of additional firms into the industry it is naturally to be expected that it will do what it can to give that firm a chance to survive. An infant-firm argument often leads the regulator to extend special protection to the new enterprise, intended to continue until that firm attains ‘maturity’ and acquires strength sufficient to enable it to fend for itself. The privatized firm -- the earlier sole incumbent -- may be required to supply services to the entrant at especially low prices, or to offer it other forms of implicit subsidy, or still other forms of protection may be provided. It is usually promised that all of these will be phased out at a suitable time, but normally no date is specified, nor is anyone told the precise circumstances when that will occur (e.g., when the entrant’s sales reach X percent of the industry’s). No one is told whether subsidies will all be removed at once or whether it will be done gradually, and if the latter, at what rate. All this imposes unnecessary uncertainty not only upon the privatized firm, but also upon its new rival. And as indicated, much of that cost will be borne by consumers.

(v) Pointless restriction upon management’s freedom of decision

The large privatized firm is predictably distrusted by the regulator. Even when the latter adopts rules ostensibly designed to reduce restrictions upon management, steps will often be taken to curtail that freedom or eliminate it altogether. For example, regulation has in recent years made use of floors and ceilings upon prices, with the bounding magnitudes based in a rough and ready way upon economic analysis. This suggests that once such limits are determined the privatized firm will be left free to select the intermediate price that best suits its interests and changing

market conditions. However, regulators are often shocked that management should be given such unrestricted license. They seek to narrow the firm's options further, or require a waiting period before the proposed prices can be put into effect, or subject the prices adopted to ex-post review and penalties.

There are at least two costly consequences. First, it restores a feature of traditional regulation which has long been criticized: the delays it imposes on the decisions of the regulated firms and the resulting lag in adaptation of its decisions to evolving market conditions. Second, it all but removes the influence of the market upon the price-setting process, ensuring that privatization does not serve as a step toward adoption of the market mechanism as the prime guide of economic activity.

(vi) Adoption of discredited regulatory criteria

The privatized firms often find themselves regulated with the aid of accounting conventions, notably *fully-distributed (or allocated) costs*, that are universally admitted to be arbitrary, that only by happenstance will bear the slightest resemblance to the costs economic analysis has shown to be pertinent to economically efficient decisions, that undermine incentives for innovation, and that often serve as protectionist devices undermining true efficiency.

Fully-allocated costs are the accountants' attempt to provide figures resembling average costs for each of the firm's products in a multi-product enterprise. The results are always arbitrary because there are typically substantial costs fixed and common to two (or more) company products, A and B, and there is no way based on the pertinent facts to determine what share of those costs is properly attributable to A, and what share to B. The result is that speciously associated criteria (for example, the value of the output of A relative to that of B, or the relative weights of the products) are used to apportion those costs arbitrarily.

Because the resulting figures generally bear no resemblance to marginal costs or any other real and pertinent cost figure, prices based by the regulator on fully distributed costs will generally lead to outputs, sales, investment levels, and so on, that have no resemblance to those required for economic efficiency. Because the fully-distributed costs are intended as substitutes for 'average costs', product by product, regulators are given to believe that prices set at those levels will ensure the firm's solvency. But those prices are set with absolutely no consideration of the different demand conditions faced by the various products, and those prices will therefore often prove uncompensatory. Because such prices are 'cost plus' in character, they eliminate any incentive for process innovation or other cost cutting efforts. Moreover, because of their arbitrary character, the fully-distributed cost figures lend themselves to manipulation and they have often been used in litigation before regulatory agencies by firms determined to protect themselves from the setting of low prices by more-efficient rivals. All this was experienced in the US for many decades in the regulatory arena. And much of this is now being reexperienced by the newly privatized enterprises.

1.6 What Is To Be Done?

This, then, is the story of the perils that beset the regulatory process so often following in the wake of privatization. Such regulation, as we have seen, is generally not unjustified, though it should, of course, be avoided wherever and whenever the market is capable of doing the job. The problem, as we have seen, is that there are so many traps that can ensnare the regulatory process and misdirect it in ways that appear calculated to enhance the welfare of the public, while in reality they damage it severely.

The question, then, is whether there are ways of doing it better -- of carrying out the process in a manner that delivers the benefits of privatization and not merely their promise. Economic analysis provides the basis for what appear to be rational answers. These answers have been explored analytically and, very recently, have been introduced into practice in a number of parts of the globe. Thus, these answers are an example of combined theory and practice in economics. The next section will describe some of those approaches and explain their logic.

2. TOWARD MORE-PROMISING MECHANISMS FOR REGULATION OF PRIVATIZED FIRMS

Out of the discussion that has accompanied the period of deregulation in the US--the period since the mid-1970s -- and the subsequent experience, there has emerged a new body of principles for the guidance of economic regulation. These principles are designed to minimize interference with economic efficiency, to expand the role of the market as far as seems advisable in areas of the economy where the strength of competitive forces is suspected of being inadequate and, incidentally, to reduce the amount of litigation with its propensity to waste large quantities of resources and, what is arguably even more serious, to inhibit the spirit of enterprise of those who run it. There is reason to believe that what may be dubbed ‘the new regulatory principles’ have, at least so far, largely lived up to their promise, and lightened the burden of regulation significantly, while contributing to efficiency. It would appear that those who regulate the newly privatized industries can profit from a study of those principles.

Those principles can be summarized rather briefly.[5] I shall, therefore, start out by doing so and then I will go, in greater detail, into two of the issues that are rather more subtle -- the regulation of profit in ways that preserve the incentive for innovation, and rules for pricing of bottleneck services (services supplied by only a single firm both to itself and to rivals who use that bottleneck service as a critical part of the process of supply of a final product in which all these enterprises com-

[5] For a fuller discussion, see Baumol, W.J. and J.G. Sidak (1994).

pete). In the current section, methods for dealing with these issues will be described and their benefits explained.

2.1 The General Principle Underlying the New Regulatory Approach

The underlying premise of the new approach is that the sole purpose of economic regulation is to facilitate and encourage effective competition where that is feasible, and to provide an effective substitute for competition where that is not possible, at least for some substantial period. In a later section appropriate means for the encouragement of competition will also be considered. Here I focus on the latter regulatory task, that of serving in *loco competitio* -- i.e., as a substitute for the absent competitive forces, no more and no less. In other words, the underlying premise of the new regulation is that where competition is effective it can do a better job of protecting and promoting the public interest than any government agency. Therefore, where, and only where, competition is either absent or too feeble to do the job, it is appropriate for the regulator to step in. But in doing so, the regulator's obligation is severely limited. It is to supply as near a substitute for the missing ingredient as can be devised, that is, to determine means to elicit the business behavior that effective competition would have enforced if only it had been present.

If it is agreed that this is the proper task of the regulators, then two things follow at once. First, it is their obligation, in markets where the strength of competition is deemed inadequate, to constrain regulated firms to adopt only such decisions and act only in such ways that effective competition would permit. That is, the regulators must permit firms to choose only among courses that would be open to those enterprises if, contrary to fact, the markets had been effectively competitive. Second, it follows from our specification of the regulators' underlying role that *they must not constrain the firms under their jurisdiction any further than this*, that is, the regulators must accept a self-denying ordinance obligating them never to prevent managements from any action that they could have carried out in an effectively competitive market. These observations lay out the range of freedom of choice that the underlying principle requires firms to be guaranteed, and it also indicates in general terms the appropriate limits to be placed on this freedom.

The task of the regulator, then, consists of two parts. First, it must determine which choices competitive markets do, and which they do not leave open to firms and, second, it must adopt procedures to ensure that the firms will act in a manner consistent with the competitive standard.

The literature of economics provides considerable help in carrying out the first of these tasks, for it contains very substantial discussions of the behavior of competitive industries (as well as analyses of the virtues of competition). Here, one *caveat* applies. The industries containing newly privatized firms will often be characterized by scale economies, at least up to some rather substantial level of output. Hence, a large multiplicity of firms probably will neither be feasible nor desirable, and marginal cost pricing -- the type of pricing that characterizes the

theoretical regime of perfect competition will very likely be incompatible with solvency of the firms. Thus, the competition that serves as the standard for regulators here is not the model of perfect competition. Rather, the equally theoretical concept of *perfect contestability*, defined as the state of a market in which there is totally costless and unimpeded freedom of entry and exit, must serve as the model. That is because, unlike perfect competition, the definition of perfect contestability makes clear that it is compatible with the presence of scale economies and the existence of only a small number of firms or even only one firm, which is, hardly accidentally the state of affairs in the preponderance of regulated industries, particularly those that are newly privatized. Moreover, contestability theory does derive, in considerable detail, the requirements of economic efficiency in such circumstances (see Baumol, Panzar and Willig ,1988). That is, the contestable-market analysis does offer the regulator guidance on the regulatory rules for the behavior of firms that will provide the economy with the benefits of effective competition.

2.2 Some of the New Rules for Economic Regulation

We can now summarize very briefly the rules and principles that emerge from the analysis, providing even a short explanation only where it seems necessary at this point. Later in the section two of the rules (6 and 7) will be gone over in greater detail.

1. In any market where there is evidence that competition is sufficiently powerful to protect the public interest, regulators should refrain from intervention.

2. In markets in which adequate competition (rather than the mere presence of a multiplicity of non-competing firms) can be stimulated, that should be done.

3. Prices should not be permitted in the long run to exceed the levels that in a perfectly competitive market would make entry profitable, entry that would subsequently drive those prices back down. These price ceilings are referred to in the literature as the ‘stand-alone costs’ of any product or combination of products.

4. Prices should not be permitted to go below those that would be viable for any substantial period in a competitive or contestable market. This generally means that those prices should not fall short of the marginal cost of any product or the per unit incremental cost of the entire output of any homogeneous product.

5. Because in a contestable market one may encounter prices close to the stand-alone cost ceiling or the marginal-average incremental cost floor, the firm should be left free to adopt any price within these limits, adjusting that choice to current demand conditions in accord with the judgment of management.

6. *Price caps should be substituted for fixed ceilings on total earnings or for a fixed ceiling on rate of return on investment.* That is, the regulator should set upper limits on prices, rather than profits, so that the firm will be offered the incentive to cut costs. Price ceilings are not to be adjusted downward immediately to correspond to any reduction in costs the regulated firm is able to achieve. Rather, in accord with the Schumpeterian model of the market's incentives for innovation and enhanced efficiency, price ceilings are to be unchanged for substantial periods, except for a built-in inflation escalator that automatically increases the ceiling in accord with some standard price index such as the CPI, *after subtraction of some number corresponding to the industry's past record of rate of reduction of cost per unit of output per year*. Thus, an industry with an average record of productivity growth of 2 percent per year would, in a year when the CPI grew 6 percent, find its price ceilings increased by 4 percent, so that it would earn profits exceeding the competitive level if and only if it managed to exceed its past 2 percent productivity growth record. During a grace period of several years the firm will be able to keep those profits as its reward for innovation. But, just as competition ensures in an unregulated Schumpeterian world, prices will ultimately be adjusted to eliminate the enhanced profits so that, thereafter, the benefits are passed on to consumers.

7. When inputs are supplied by a regulated firm, both to itself as a component of one of its final products, X, and to a competitor producer of X, then the regulated firm should charge the rival the same price for that input that the former implicitly charges to itself. This rule is called 'the parity principle', 'the optimal input-pricing rule' or 'the profit-imputation rule'. The price of the input should equal the (incremental) cost entailed in supplying it, as usual in a competitive or contestable market, *including any profit the regulated firm forgoes by the sale of a unit of input to its rival because that permits the rival to take away some sales of final product X*. Thus the price of the input to a rival should include all of the profit contribution the regulated firm obtains from the sale of a unit of final product X.

It should be noted that these rules lead to regulatory behavior very different from that often encountered by privatized firms. Characteristically, in practice less has been left to the control of the market and less freedom has been given to management. Fully-distributed cost is often used, sometimes as a price floor, sometimes as a ceiling and sometimes as the imposed price. Input prices, rather than following the parity principle, are often set so as to pass on part of the regulated firm's profits from the final product to its competitor that purchases the input from it.

Note that many of these rules are counterintuitive to the layman. For example, the parity principle (rule 7) requires the price of a widget input, whose direct incremental cost constitutes only 2 percent of the cost of final product X, should nevertheless compensate the widget maker firm for *100 percent* of the profit it forgoes from the sales of X as a result of its supply of widgets to competitors.

It seems clear that all of these procedures, as now carried out, offer considerable room for improvement from the point of view of economic efficiency and utilization of the market mechanism. For, as in much of economics, common sense is a dangerous guide to decision makers engaged in the privatization process. Those decision makers consequently can learn much from study of the pertinent principles of economic analysis. For this reason, I turn next to a discussion in greater detail of the last two of the preceding recommendations. These relate to price caps and to the parity-pricing rule, neither of which is self-evident in its logic.

2.3 Dynamic Efficiency, Price Caps and Market-Based Pricing

One of the most significant criticisms of the old regime of rate-base, rate-of-return regulation was the very limited incentives that it provided for innovation and other measures to improve product quality and enhance the growth of productivity. The ceiling on the regulated firm's rate of return deprived it of such an incentive because it prevented any profit reward from being earned by a firm that had a superior performance in terms of growth in productivity or product quality.

This is in marked contrast to the profit experience of firms in unregulated competitive markets. There, the firm whose productivity and quality performance is below the norm can expect to suffer financial losses. Similarly, the company whose performance in these areas is outstanding can expect substantial compensation in terms of the additional profits it earns by its superiority over its rivals. Thus, if regulation is to undertake a program that restricts pricing on the basis of the market model, then the regulatory rules must make provision for similar incentives for improvement of the regulated firm’s performance.

It is for this purpose that *price caps* (to be defined presently) were invented as an instrument of regulation. They are designed to constrain the pricing that a firm with some degree of market power can adopt, thereby protecting the interests of consumers. At the same time they offer management wide freedom of decision making, permitting the firm to adopt any prices that fully-effective competition would not preclude. Price caps are often also designed to extend freedom of decisions on the prices of individual products of the firm by setting a ceiling only over the *average prices* of groups of related products rather than constraining the price of each product individually. Finally, and most important for dynamic efficiency, they provide automatic incentives for the effort to increase productivity that are built into the rules and that require a minimum of human judgement and intervention to achieve their purpose.

This last point, the automatic character of the incentives provided by price caps, needs to be stressed and explained. Attempts to offer rewards for superior performance under earlier regulatory regimes generally ran into a difficulty that often proved insuperable. Measurement of the efficiency of operation of a firm is exceedingly difficult to carry out or even to define. It cannot be done in terms of costs alone because costs are often affected profoundly by developments outside the in-

dustry and often the firm can reduce costs simply by cutting product quality. Moreover, product quality itself is difficult to measure, particularly when it entails a variety of attributes of the good or service supplied. For example, the quality of passenger transportation depends on time taken in the journey on average, frequency of late arrival, cleanliness of the vehicle, seat comfort and a variety of other attributes that are hard to measure and harder to aggregate into a single quality index. As a result, regulators generally failed in their attempts to rate the regulated firms in terms of efficiency and quality of performance and to modify their financial returns accordingly. As will be shown presently, the price cap approach obviates the need to undertake any direct measurement of the performance of the regulated firm and to calibrate a suitable reward or deterrent for differences in performance among firms. Thus, price caps serve as a practical means to incorporate the market-based incentives for productivity improvement into the price regulation process. Like the free market, they provide and administer appropriate rewards and penalties without detailed regulatory intervention or oversight.

What, precisely, is a price cap and where is it currently used? A price cap is a moving ceiling over some price or over the average of some group of prices that is automatically adjusted for inflation in accord with some fixed formula, *but in which the ceiling is adjusted upward at a rate some amount lower than the actual rate of inflation.* The difference between the rate at which the ceiling increases and the actual rate of inflation is intended as an estimate of the industry's normal productivity growth, which obviously can serve as an offset to the rising price of fuel, labor and other inputs. I will presently discuss the mechanism of this adjustment and its purpose a bit more fully.

Price caps are increasingly being used by regulatory agencies as a means to move toward 'light-handed regulation' while continuing to protect the public interest against abuse of market power. This approach is now employed in the United States by a number of agencies including the Federal Communications Commission in its regulation of long-distance charges, in Great Britain in regulation of airport fees and in a number of other uses. It is also widely being considered for adoption in arenas where the rules of regulation are under active review.

Price caps are intended to serve a number of purposes:

(i) Price caps as proxy for competitive constraints

Most obviously, price caps are designed to prevent firms from charging prices above competitive levels, that is, prices that can yield monopoly profits. This means that, properly designed, price caps should be set so as to prevent the adoption of any price that could not prevail if the market was effectively competitive, but it should not preclude any price that could be sustained if the market were effectively competitive. Thus, price caps should be employed in imperfectly competitive markets as a proxy for competition, and should restrain business decisions

neither more nor less than market forces would if competition were fully effective. This also implies that price caps have no purpose and constitute costly and excessive intervention in any market where competition is patently effective.

(ii) Price caps as instrument of deregulation

It is clear that price caps give the managements of regulated firms far greater freedom of decision making than was available to them under more traditional regulatory regimes. The firm is faced with a maximum figure that serves as a price ceiling (often supplemented by a price floor) that sets a limit upon the prices that can be adopted. However, so long as that limit is not violated, management is left free to select any level of price it considers appropriate, and to change that level without delay in response to evolving market conditions. Management's pricing freedom is obviously extended if the ceiling constrains only the average price of a group of products rather than being applied individual product by individual product. Since it is generally assumed that management is in a better position than a regulator to determine appropriate prices and to keep track of and respond suitably to changes in market conditions, there is some presumption that the prices that emerge under price caps will promote economic efficiency. The elimination of the delays in the process of price adjustment that characterize a traditional regulatory mechanism probably also contributes to economic efficiency by freeing the firm from the necessity of living with prices that may have been appropriate in yesterday's circumstances but that today are an impediment to economic efficiency.

(iii) Price caps as incentive for productivity growth

Price caps also force regulated firms to pursue increased productivity relentlessly, for, under a price cap regime the firm that falls behind in the productivity race is automatically condemned to experience losses, while only the firm that beats the productivity norm can earn a return greater than the normal competitive earnings rate. In this way, the price cap acts just as pricing in unregulated competitive markets does, automatically rewarding firms that are successful in their productivity programs and automatically punishing those that fall behind.

The way that price caps do this is straightforward. By raising price ceilings at a rate equal to the rate of inflation of industry input prices, *minus the normal productivity growth of the industry*, price caps make it possible for a regulated firm to continue to earn a profit equal to the competitive rate if and only if its current productivity performance just matches the norm. If input prices are rising 6 percent per year and average industry productivity growth offsets the resulting cost increase at a rate of 2 percent a year, then the firm that matches the industry average will experience a 4 percent net increase in cost. In these circumstances the price cap will also grow at 6 - 2 = 4 percent per year, so that this regulated firm, with its

average performance, will be able to match its prices to the 4 percent increase in its costs. However, a firm with a superior productivity performance, say, one that achieves a 2.5 percent growth, will experience a cost increase of 6 - 2.5 = 3.5 percent, which is less than the 4 percent rise in the formula-driven price cap. The difference can accrue to this productivity-achieving firm in the form of additional profit that serves as a reward for its contribution to the economy's dynamic efficiency. For exactly the same reason, a firm that is a productivity laggard in this industry will experience cost increases more rapid than that of the price cap and so such a firm will undergo a corresponding loss of earnings.

In this way, a price cap puts the regulated firm under constant and unrelenting pressure to beat the norm in its productivity performance and never to fall behind the norm. The result is a substitute market incentive for each firm to work determinedly for improvement in economic efficiency. It is this incentive that traditional regulation failed to provide.

We conclude that, to the extent that regulation of prices of the privatized firm is deemed to be appropriate, there are strong reasons for employing the price-cap approach as the regulatory instrument. This approach appropriately limits the scope of regulatory intervention, provides maximal freedom for managerial decision making, can effectively prevent exercise of market power and provides a strong incentive for dynamic efficiency.

2.4 Parity Pricing of Bottleneck Services

I turn, next, to the parity-pricing rule, a second regulatory principle grounded in economic analysis that requires some careful explanation. A common issue has recently come to the forefront in a number of regulated industries, notably in the supply of electricity, telecommunications and railroad transportation. The case of electricity, as an example, will bring out the nature of the issue and make clear its importance. The activities of the suppliers of electricity can be described in two categories: generation, or the act of producing the electric power, and transmission, the transportation of that power so that it can be used by its final consumer. Until very recently both of these services were generally supplied by monopoly providers. In any given geographic area (with exceptions that are irrelevant for our analysis) there was only a single firm that carried out both generation and transmission. Recently, there has been a change in the regulatory rules, and the regulatory authorities have begun to encourage the entry of competitors into the generation process. In transmission, however, competition is still deemed to be impractical because of economies of scale, that is, because a second supplier of transmission would simply have to replicate the facilities of the current transmitter, thereby

wasting resources for society and imposing an impossible cost handicap upon the would-be competitor.

The problem in having one firm continue to supply both generation and transmission, while its rivals supply only generation is that the latter are forced to purchase transmission services from the former, who is the only supplier of transmission in the area. The problem is that the transmission monopolist, if unrestrained, can charge a transmission price so high that it threatens to drive its rivals out of the generation business.[6] In the discussion that follows, it will be helpful to think of electricity generators as firms that manufacture electricity, and then purchase transmission service in order to sell the final product, the electric power they created, to the power consumer. Thus, the final product, delivered electricity, is sold by an independent firm using two inputs: generation, which it supplies itself, and transmission, which it purchases from the owner of the bottleneck transmission facilities.

(i) Why the issue is difficult?

If the provider of bottleneck services such as electricity service were not also a supplier of the final product (electric power sold to consumers) it would be easy to determine a price for access to the bottleneck input that constitutes no competitive impediment to one final-product provider vis-a-vis any other. The obvious solution is to require the supplier of the bottleneck to charge exactly the same price to all final-product providers *that compete with one another*. By avoiding differential pricing in the sale of bottleneck services to rival final-product providers they are left free to compete for customers strictly on the merits. Specifically, since any two rival suppliers of the final product, then, are paying the same price for the bottleneck component of, or input of, the final product, the firm that can provide the non-bottleneck portion of the final product more cheaply can always afford to undercut its rival in the long run. Indeed, since we can think of the supply of the final product (delivered electricity) as composed of two components, the bottleneck service (transmission) and the remainder (generation), *when both firms pay the same price for bottleneck services, the firm with the lower non-bottleneck cost can afford to undercut its competitor precisely by the amount of the difference in their costs*. If firm A can provide the non-bottleneck part of the final product at a cost that is lower than B's by X cents per unit of final product, A's charge for that product can also be X cents per unit cheaper than B's.

Reality, however, does not provide such an easy solution to the handicapping-avoidance issue. The difficulty stems from the fact that bottleneck service is sold by firms that supply final product of their own. In reality, the electric utilities that own the transmission facilities provide transmission service to the independent

[6] There are, actually, good reasons why a transmission monopolist may find it in its self-interest to refrain from doing so. But that is a matter beyond the scope of our discussion here.

generation-supplying companies as well as providing it to themselves. In other words, the utilities may sell transmission service both to themselves in their role as sellers of electric power to electricity users, as well as to their rivals in the sale of electricity. This immediately raises the suspicion that the utilities will be tempted to favor themselves over their rivals in the supply of transmission service, but the problems go deeper than that. For, it is not even obvious at what transmission price the utility would be treating every electricity supplier equally.

The obvious solution that, unfortunately, is not very helpful, is to require the utility to charge itself for access service exactly the same price that it charges all rival electricity providers. This statement is correct, and does underlie the parity-pricing rule. However, the price that the utility really charges itself for access is far from obvious. Such a price may be specified in the firm's accounting records, but that price is *really* an artificial and arbitrary number that tells us nothing about what the utility really gives up financially (that is, what it really costs that firm) when it supplies transmission service to itself. (It is convenient here to think of the utility as a firm that is composed of two divisions -- one division generates electricity, the other transmits it.) So, the accounting figure that purports to be the interdivision electricity transmission price can be raised or lowered arbitrarily without any financial consequence for the utility. Such a rise in the accounting access price simply means that a correspondingly smaller profit contribution is credited to the electricity generation division of the firm in the company's books, but that must be offset precisely by an equal increase in the profit imputed to the company’s transmission division. In other words, if the accounting-access price were used as the standard for parity pricing, the utility could, if it wished, adopt any figure it liked for the purpose, to support a claim that it was charging itself for transmission exactly the same price it is charging competing toll providers. Clearly, the accounting transmission price cannot do the job. It is necessary to search further in order to determine what price the utility is really charging itself for the bottleneck services it provides to itself.

(ii) The parity-principle formula for bottleneck-service pricing

The analysis underlying the parity principle solves this problem. It tells us that the price that the utility charges itself for access of one kilowatt-hour (kwh) of transmission is simply the price the firm charges to the final customer per kwh of electricity, minus the incremental cost to the utility of the generation of that electricity (the cost caused to it by an additional kwh of generation), including in that cost the competitive return on the incremental capital required for the purpose. The parity principle tells us that this subtraction reveals the true price that the utility implicitly charges itself for transmission and it is, consequently, the price at which competing generation providers should be entitled to purchase transmission from the utility.

There are two ways to confirm that this is the correct valuation of the transmission price that the utility implicitly charges to itself. The first is purely intuitive. It notes that the final product price consists of two portions, the part accounted for by transmission and the part attributable to the remainder, i.e., to generation. Therefore, if one subtracts out from the utility's final product price the portion attributable to generation, then the remainder should be the part that corresponds to the price of transmission. However, the generation activities in the circumstances under consideration are supplied under competitive conditions, and the normal price of such a product should cover its incremental cost including the competitive rate of return (i.e., the fair rate of return) to the required investment. This is precisely the amount that is deducted in the parity principle calculation to determine the price the utility is paying itself for transmission.

The discussion of the preceding paragraph is just impressionistic. However, there is a logic-tight way of showing that the formula that has just been described does correctly give us the number that is required. The proof is rigorous but indirect, and the full analysis must be carried out mathematically, so that I have confined that proof to a footnote. Its logic is, however, not difficult to understand. Earlier it was noted that if and only if an independent supplier of transmission services is charging the same transmission price to two independent sellers of the electricity they generate, then the difference in the prices at which those firms can afford to sell a kwh of electricity to final consumers will be exactly equal to the difference in their (incremental) generation costs. This obvious proposition can be extended directly to a utility firm that is a provider of both generation and transmission. For it is obvious that the utility will then be selling transmission service to itself at the same price that it sells that service to a rival electricity provider if and only if at that transmission price the rival can afford to sell electric power at a price that differs from the utility's by precisely the amount that the rival's incremental generation cost differs from the utility's. If the generation cost of the competitor is X dollars per kwh lower than that of the owner of the transmission facilities, then both are paying the same transmission price if the rival can afford to provide electric power to consumers exactly X dollars cheaper than the utility can. The opposite should obviously hold if the utility is the more economical generator. One then proceeds by determining the transmission price that permits the final product prices of the two firms to differ in just that way. And as is proved in the footnote, this will be true if the price of transmission to the competitor is set equal to the utility's final product price, minus the utility's incremental cost of the generation portion of electricity-supply processor. For then the competitor can sell its own product at a price equal to the transmission price plus its own incremental generation cost (plus the competitive return on its capital). Hence, the rival's lowest viable final product price will equal its own incremental generation cost plus the utility's final product price minus the utility's incremental generation cost. In other words, when transmission is priced to the rival at the utility's final-product price minus its incremental generation cost it must be true that the rival's minimum elec-

tric-power price = the utility's toll price plus the rival's incremental generation cost minus the utility's incremental generation cost.

Thus, parity pricing of the bottleneck service permits the rival to undercut the final product price of the bottleneck owner only if the rival is the more economical supplier of the non-bottleneck portion of the final product, and it then permits the bottleneck owner's final product (delivered electric power) price to be undercut precisely by the amount of the rival's incremental cost advantage over the bottleneck owner in the non-bottleneck portion of the final product.[7] This shows why the

[7] All of this can be proved systematically in the following manner:

Proposition: The parity price for a bottleneck service is both necessary and sufficient in order for the playing field to be level, i.e., for the maximum difference between the remunerative prices of the perfect-substitute final-products of the two firms, the bottleneck-service provider (B) and its final-product competitor (C), to be exactly equal to the difference in their incremental costs for the non-bottleneck portions of their competing final-product supply.

Notation:

$P_{finalprod,b}$ = the bottleneck owner's *given* price of final product (say, per kwh).
$minP_{finalprod,c}$ = the competitor's minimum viable price of final product.
$P_{botservice}$ = bottleneck-service price per unit of product (e.g., per kwh).
$IC_{non\text{-}botserv,b}$ = the incremental non-bottleneck service (generation) cost to the bottleneck owner (the utility) per unit of final product.
$IC_{non\text{-}botserv,c}$ = the corresponding figure for the competitor.

Proof: By definition,

(1) $minP_{finalprod,c} = P_{botservice} + IC_{non\text{-}botserv,c}$.

The level playing field is defined by

(2) $minP_{finalprod,c} = P_{finalprodl,b} - IC_{non\text{-}botserv,b} + IC_{non\text{-}botserv,c}$.

That is, the lowest compensatory price the competitor can charge should differ from the bottleneck owner's exactly by the amount that the former's non-bottleneck costs fall short of the latter's. Comparing the two equations, we see at once that the level playing-field condition (2) will be satisfied if and only if

(3) $P_{botservice} = P_{finalprod,b} - IC_{non\text{-}botserv,b}$.

But this is the parity pricing formula. Thus, parity pricing is both *necessary and sufficient* for a level playing field. **QED**.

parity principle has been said to level the playing field between the bottleneck owner and competing final-product suppliers.[8]

We conclude that the parity pricing rule does provide a formula for the determination of bottleneck-service price that is straightforward to use, and that this formula does, indeed, ensure that the playing field for the bottleneck owner and its rivals in final-product supply is level.

2.5 Can Competition Really be Encouraged by the Regulator?

Despite the frequency with which, in its earlier life, the privatized firm held a monopoly position to which government raised no objection, once the enterprise is privatized it is not unusual for the regulators in charge to devote themselves to the introduction of competition into the market. Once the entry of a second or a third firm has been facilitated or even elicited by the regulators, they are undoubtedly convinced that they have thereby contributed to competition, even in cases where this has occurred through the imposition of a territorial-division arrangement in which the entrant is effectively shielded from effective competitive pressures. Once entry has occurred, the regulators' dedication to competition is adjoined to their natural predisposition to ensure the survival of every enterprise under their jurisdiction, no matter how inefficient and costly to the public.

Though these observations are tinged with a critical tone, it must be conceded that those regulators do have a very valid point. If it is ever to be appropriate to free the privatized firms from regulation and thrust them, unfettered, into the free market, competition must somehow be brought into the arena. One is led naturally to ask, then, is this really possible? More specifically, are the cases in which true competition is a realistic possibility rare exceptions, or is this state of affairs relatively common in the industries where privatization has occurred? Finally, in cases where one cannot be confident that competition will evolve by itself, or there are good grounds to fear that it will do so on too modest a scale or at too slow a pace, what measures are suitable for its facilitation and encouragement? Here, I have little more than impressions to offer, because the evidence on these matters has hardly been explored. Thus, unlike the earlier parts of this section, this section can claim little support for its conclusions from formal economic analysis, though they do rest, to some degree, on experience in the arena.

One cannot dispute the standard conclusion that in fields in which scale economies are strong, universal and prevalent through all of the relevant ranges of output

[8] This is misleading in one important respect, however. Even if the price of bottleneck service is the same to bottleneck owner and its competitor, the price can be higher to both of them than is justifiable. Such monopoly pricing of the bottleneck services will lead to a rise in the price of the final product, reducing the demand for that item and thereby harming consumers. Such a restriction in demand is also likely to affect one of the firms more than it does another. In other words, the public interest requires a parity-pricing rule to be supplemented by others that prevent monopoly overcharge for the bottleneck service.

quantities, monopoly is ‘natural’. That is, in these circumstances, it is unlikely that a multiplicity of firms will be able to survive, and it is, moreover, probable that their survival is undesirable. The reasons are well known, and hardly need repetition here. The small-scale entrant has poor prospects for survival because its very size is the source of costs higher than the incumbent's. At the same time, survival of the entrant means that its higher cost must be borne by consumers, and that can hardly be considered to constitute a benefit to them.

Casual observation suggests, however, that such cases may be rare. The evidence does indicate that scale economies do arise in a considerable number of industries, but the evidence relates only to a narrow range of output levels beyond the current outputs. In the absence of experience in the outer reaches of the pertinent outputs the econometric studies, quite understandably, provide little evidence about the full range of outputs over which economies of scale and scope prevail, and offer no indication of the points at which they are exhausted. But one does notice that while oligopoly is a fairly common phenomenon, monopoly that has not been imposed by government seems rather rare. Indeed, outside a few public utilities, it is difficult to think of examples.

This would seem consistent with what empirical studies indicate -- that in a number of industries scale economies are substantial, that they prevail over a considerable range of outputs, but that beyond some output levels they are replaced by approximately constant returns to scale that themselves hold over significant output ranges. This is the equivalent of the observation for the (theoretical) single-product firm that the average cost curve is more realistically taken to be flat-bottomed rather than U-shaped.

The hypothesis implicit in the preceding paragraph, then, is that among privatized industries the multiproduct equivalent of flat-bottomed average cost curves is common, but that the flat-bottomed range follows only after a wide range of substantial scale economies. If this hypothesis proves correct, it has implications for the prospects for competition in the privatized industries that are not necessarily in conflict with the views of regulators described earlier in this section.

First, it follows that at least oligopolistic competition is very possible in these industries, and that such competition can be expected to endure. Second, it suggests that successful entry is likely to require the assembly of large quantities of capital and other resources, because only firms of considerable scale will be able to compete successfully. Thus, initial entry is apt to be difficult. Third, it suggests, because of the considerable region of constant returns to scale, that the successful firms in the industry need not be of similar size, and that relatively large and relatively small firms may be able to coexist. Finally, it follows that even the smallest firms in a long-run equilibrium in such an industry may prove to be very effective competitors, able to exert a strong constraining influence upon the pricing of the larger enterprises in the same market.

In those privatized industries where this is the state of affairs it follows that at least some degree of competition can be achieved, and that the eventual presence of

a number of competing firms can be hoped for. If collusion can be prevented or is inherently unlikely, the prospect (even with a small number of rivals) is that the firms will compete vigorously and effectively, though such competition may make use of strategic courses of action that yield public benefits short of those to be expected from a perfectly competitive or contestable industry. Yet, particularly where entry requires substantial sunk investments, one cannot be fully confident that one will experience the establishment of the new firms requisite for transformation of the market into one that is highly competitive or at least highly rivalrous. Experience indicates that such entry can in fact take place. For example, the successful establishment of a number of long-distance carriers in US telecommunications so soon after entry was permitted suggests that private initiative can at least sometimes suffice to carry out the task.

Still, there are at least three reasons why it may not happen, even in arenas in which it is called for by the public interest. All of them relate to the impossibility of successful entry on a small scale in our scenario.

First, in the scenario under discussion successful entry requires the establishment, in one initial step, of a firm already sufficiently large to be able to compete effectively in the scale-economies industry. Consequently, the very size of the sunk investment required for successful entry can effectively inhibit private initiative from taking advantage of socially-promising opportunities. The need for such sunk investment can be exacerbated by lack of experience in the field by prospective entrants. These prospective entrants may lack the skills required for success in the arena, skills that can be attained only through learning-by-doing. Where substantial experience is required for viability of the firm, even a prospective entrant that, given the opportunity to survive for a sufficient interval of time, promises to become fully competitive and successful, may well decline to undertake the requisite outlays, waiting time and risks, all of which can be translated as investments awaiting an uncertain return in the relatively distant future.

Second, entry may be desirable to society even though it is unable to attract sufficient private investment, because the general risk to the private investor is considerably greater than the risk to society. There are many reasons why this may be so. Most notably, there is the possibility that the new enterprise may become insolvent and be lost altogether to the original investors, but that it will then undergo reorganization under new ownership, and continue to yield benefits to the economy. Where this scenario is a possibility it follows that, other things being equal, the expected payoff to society will be greater than that to private investors, so that the undertaking may be worthwhile socially, but not to any private group. A particular variant of this problem arises when there is any likelihood of successful strategic countermoves to entry by the earlier incumbent. If entry requires sunk outlays, that possibility will increase the risk facing the new firm and is likely to raise the cost of funds to the entrant, though its assets may continue to serve the society even if the strategy of the predecessor firm is successful.

Finally, there is the possibility that externalities, perhaps the most usual cause of market failure, will be present -- i.e., that the act of entry will yield socially beneficial spillovers from which the investors do not profit. The most obvious form that this can take occurs when entry frees the economy from the costly burdens of regulation in this arena. If the presence of a multiplicity of firms is deemed to render regulation redundant in the industry in question then that may constitute a substantial benefit both to the earlier incumbent and to the body of customers, benefits that are not reflected in the earnings of the entrant. Once again, it follows that entry may be unprofitable even if it is socially beneficial.[9]

All of these observations can constitute justification for some public sector intervention to facilitate the entry of new firms and encourage their growth. In general, the externalities problem aside, it is to be expected that the danger to survival against which they should be protected will be temporary, for otherwise there is a real question as to whether their presence will ever constitute a significant contribution to competition, or any contribution at all. The issue, then, is a variation on the infant industry theme. Enough has been written on this subject to make redundant any further discussion of the validity of the infant industry argument and its implications. There are, however, some observations that may be illuminating that grow out of experience of attempts by regulators of privatized firms to provide protection to infant entrants.

The problem is that there exist forms of protection of infant firms that are extremely and unnecessarily costly to society. For example, rules that force the incumbent firm's prices substantially higher than are called for by its costs can make life easier for the entrant, but they also are sources of inefficiency because they reduce the incentive for the entrant to invest in growth in its productivity or, more generally, to learn how to comport itself effectively in a real competitive battle. Similarly, the imposition of an arrangement that assigns portions of the market exclusively to the entrant, clearly impedes efficiency, beside contributing little or nothing to competition. Yet both of these devices are apt to recommend themselves to the regulatory agency in its well-intentioned attempt to foster competition.

The discussion in earlier paragraphs of this section of the reasons the market may fail to elicit all the entry that is socially desirable also suggests more-efficient ways to provide public sector encouragement to the entrants. For knowledge of the source of the difficulty implies much about its best remedy. As we have seen, the difficulty seems to have two primary sources: capital rationing, which can be interpreted as an excessive private cost of the capital required for entry, and externalities deriving from the presence of the entrant. But externalities and differences

[9] Entry by private enterprise can also be deterred by the absence of nearby ancillary industries, whose presence would have reduced the costs of the new firm. In this case, the problem is not that the entrant provides beneficial externalities to others for which it is not compensated, but that the entrant is deprived of the external benefits that nearby ancillary industries would have provided if they had been present.

between private and social costs more generally are well understood by economists, and efficient remedies for them are described even in elementary economics textbooks. In the case under discussion what is clearly called for is a subsidy for the entrant's borrowing, in accord with the standard neoclassical analysis. Ideally, of course, it should be financed by the public treasury, but that is unlikely to be feasible politically. In practice, it has proven far easier to impose such burdens on the incumbent privatized firm for which it is hoped to elicit new competitors. Economic analysis indicates that such a required subsidy of the entrant by the incumbent has efficiency costs of its own. However, if there is no alternative, it is surely least damaging to require the incumbent to establish a capital subsidization fund for the entrant, without constraining the incumbent to employ for the purpose governmentally specified sources (for example, uneconomically high prices for particular products). The flexibility called for here is directly analogous to the flexibility permitted by effluent charges upon the emitter of pollutants, as compared to the use of direct controls as means to control pollution, with the superiority of the former approach deriving in good part from the freedom that it gives to polluting firms to seek to reduce their pollution charges by finding low-cost ways to reduce their emissions.

One final point is appropriate here. If governmental assistance to entrants is to be a temporary affair there is much to be gained if the time path of reduction of such assistance is made as clear as possible in advance, and the date at which the assistance is scheduled for elimination or the circumstances in which it will be terminated announced well in advance as a commitment of the regulator. Such precommitment offers two clear benefits. First, it eliminates the well-recognized danger that the infant firm will never be deemed to have grown up, and that its protection will be continued indefinitely. Second, it avoids the creation of unnecessary uncertainty, reducing this source of substantial cost for incumbent and entrant alike. Finally, it provides an added incentive for the entrant to prepare itself for the rigors that will be entailed in having to fend for itself in a competitive market place, and thereby encourages the entrant to invest in efficiency and strengthening of its competitive position.

2.6 Concluding Comment

This section has sought to illustrate the practical insights that economic analysis provides for regulation of the privatized firm and for effective limitation of regulatory intervention. It has shown that there exist a variety of approaches capable of reining in any substantial market power possessed by the regulated firm without, at the same time, insulating it from the forces of the market or from effective regulatory substitutes where the market is incapable of doing the job. We have seen that the regulatory measures in question can stimulate both static and dynamic efficiency. These observations clearly illustrate what practitioners can learn from economic analysis. But this is not a one-way street. The very analyses upon which the

discussion rests were themselves stimulated by the experience of economists in the workings of industry and its regulation. It is this experience that has helped to make the suggested regulatory rules practical, and their practicality is indicated by the extent to which they are being considered and actually adopted, sometimes with the consent of all the immediately-affected parties. Analysis and practice can, indeed, learn from one another.

BIBLIOGRAPHY

Baumol, William J., John C. Panzar and Robert D. Willig (1988), *Contestable Markets and the Theory of Industry Structure*, San Diego: Harcourt, Brace, Jovanovich, Revised Edition.

Baumol, William J. and J. Gregory Sidak (1994), *Toward Competition in Local Telephony*, Cambridge, Mass.: MIT Press.

Johnson, Christopher (1992), *The Economy Under Mrs. Thatcher. 1979-1990*, London: Penguin Books.