Appropriate Regulatory Technology

The Interplay of Economic and Institutional Conditions

Leroy P. Jones

Appropriate regulation means maximizing the benefits from removing market failures in relation to the costs of government intervention. Monopoly markets fail because of both allocative and cost inefficiencies. The former are measured by Harberger's little triangles and the latter by big rectangles. Regulation that mitigates allocative inefficiency while exacerbating cost inefficiency is inappropriate because the costs of intervention outweigh the benefits. In developing and formerly socialist countries the potential benefits from intervention are larger because the realm of market failure is greater, but the costs of intervention are also larger because government failures are more likely. The marginal benefits of regulation decline linearly as intervention increases, while costs rise exponentially. Therefore intervention should only attempt to control egregious allocative inefficiencies through low-cost mechanisms. The Chilean and New Zealand methods represent quite different ways of doing this and thus provide appropriate models for developing and formerly socialist countries.

In 1991 Mongolia adopted the American Telephone and Telegraph (AT&T) model for dealing with its telephone monopoly, breaking it up into twenty-three separate companies. The question is whether this transfer of regulatory technology was appropriate for a country with only 2 million people, very low telephone penetration, and only a handful of economists, accountants, lawyers, and managers trained to work in market economies, and without specific regulatory bodies or even the institutional framework for enforcing legal obligations.

Leroy P. Jones is professor of economics at Boston University. He would like to thank Mukesh Chawla for effective research assistance and Guy Beatson, Henry Ergas, Jean-Jacques Laffont, David Sappington, and Ingo Vogelsang for helpful comments.

© 1994 The International Bank for Reconstruction and Development / THE WORLD BANK 181
The example of Mongolia, albeit extreme, illustrates the general issue addressed in this paper: how to regulate market power in developing and formerly socialist countries. On the one hand, these countries have smaller economies, which exacerbates the problems of market power. On the other hand, they have weak institutions and smaller endowments of appropriate human capital, which makes dealing with these problems more difficult. The interplay of these economic and institutional conditions is the focus of this paper. The goal is to help identify regulatory institutions that are appropriate to these conditions. Or, by analogy with the effort to identify technology appropriate to the factor endowment of developing countries, the search is for regulatory technology appropriate to the factor and institutional endowments of developing countries and formerly socialist economies.

Benefits and Costs of Regulation

The benefits of regulation can be thought of as the welfare gains from correction of market failure to achieve a Pareto-efficient outcome. The costs of regulation are those that follow from government's failure to rectify matters without introducing additional distortions. Both benefits and costs need to be defined broadly. The costs include the opportunity cost of the public and private resources devoted to the regulatory process, as well as the possibility that imperfect institutions operating with imperfect information may create new perver- sions in the attempt to eliminate old ones. Similarly, market failures include not only the problems created by market power but also those associated with imperfections variously labeled as transaction, agency, contracting, or information costs.

Failures in Regulating Monopoly Pricing

To make these ideas more concrete, consider the market failures that occur in the production and management practices of a natural monopolist who is producing a nontradable good (table 1).

Although these lists are by no means comprehensive, they indicate the classes of costs and benefits associated with regulation. They also suggest that there is considerable variation in the magnitude and signs of the costs in various institutional environments. The point is simply that appropriate regulatory policy is designed to maximize the net benefits for a particular industry, country, and time. Although the list is intentionally even-handed in balancing the arguments of the left and the right, in particular cases the costs and benefits will be anything but balanced. How large might the various costs and benefits be?

Magnitudes of Net Benefits

Arguments for No Regulation. The set of regulatory options notably includes no regulation at all. The case can be and has been made that the costs of
Table 1. Market Failure versus Government Failure

<table>
<thead>
<tr>
<th>Market failure</th>
<th>Government failure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. A profit-seeking monopolist will reduce the</td>
<td>1. Government efforts to regulate monopoly reduce welfare by the opportunity costs of the</td>
</tr>
<tr>
<td>quantity offered and thus increase the price at</td>
<td>resources devoted thereto. This includes not only the costs of the regulatory body but also</td>
</tr>
<tr>
<td>which it sells. The resulting allocative ineffi-</td>
<td>the costs of the regulated firm and of the consultants hired by both sides.</td>
</tr>
<tr>
<td>ciency is the first evil of monopoly and is re-</td>
<td></td>
</tr>
<tr>
<td>flected in Harberger's little triangles (figure 1),</td>
<td></td>
</tr>
<tr>
<td>as described in any introductory economics text.</td>
<td></td>
</tr>
<tr>
<td>2. If ownership is divorced from control and</td>
<td>2. Regulators are not perfect. One consequence is that they are not fully able to accomplish the</td>
</tr>
<tr>
<td>agency problems prevent the owner from fully</td>
<td>goals of regulation. Much more dangerously, imperfect regulation may introduce new distortions.</td>
</tr>
<tr>
<td>imposing his objectives on the manager, the</td>
<td>One heavily studied manifestation of this phenomenon is the Averich, Johnson, Wellisz</td>
</tr>
<tr>
<td>manager will use his discretion in ways that</td>
<td>(AJw) effect, whereby excessively generous estimates of a “fair” rate of return lead to excess</td>
</tr>
<tr>
<td>divert some of the monopoly rents to serve his</td>
<td>investment (Kahn 1992). Under different institutional arrangements, the opposite effect can</td>
</tr>
<tr>
<td>own objectives (Williamson 1975). This may take the</td>
<td>occur, with underinvestment following from enterprise fears that government will use price</td>
</tr>
<tr>
<td>form of overspending or perquisites (activities such</td>
<td>regulation to expropriate future earnings (Levy and Spiller, in this volume). Reversing our</td>
</tr>
<tr>
<td>as entertainment and fancy means of transport that</td>
<td>field once again, there is a considerable body of literature which argues that, far from</td>
</tr>
<tr>
<td>serve corporate goals but also contribute to the</td>
<td>exploiting the regulated, the regulators go to bed with them. This regulatory capture</td>
</tr>
<tr>
<td>manager’s utility). It may take the form of</td>
<td>provides a fig-leaf of consumer-protection respectability but actually leads to high prices in</td>
</tr>
<tr>
<td>increased managerial leisure, which results in</td>
<td>the short run and protection from competitive entry in the long run. Many other examples are</td>
</tr>
<tr>
<td>less supervision of subordinates and higher costs.</td>
<td>possible.</td>
</tr>
<tr>
<td>It may take the form of reduced bargaining with</td>
<td></td>
</tr>
<tr>
<td>suppliers and labor. All these increase the</td>
<td></td>
</tr>
<tr>
<td>measured costs of the firm (the big rectangle in</td>
<td></td>
</tr>
<tr>
<td>figure 1). The magnitudes can be substantial. One</td>
<td></td>
</tr>
<tr>
<td>study of U.S. corporations finds that up to two-</td>
<td></td>
</tr>
<tr>
<td>thirds of monopoly rents are diverted to employees</td>
<td></td>
</tr>
<tr>
<td>alone (Karier 1985). I will refer to this</td>
<td></td>
</tr>
<tr>
<td>interchangeably as the second evil of monopoly,</td>
<td></td>
</tr>
<tr>
<td>the corporate governance problem, the X-efficiency</td>
<td></td>
</tr>
<tr>
<td>problem, or the rectangle problem.</td>
<td></td>
</tr>
<tr>
<td>3. While some of the behavior described above will</td>
<td>3. Government regulators have the power to grant or deny rents to firms. In some</td>
</tr>
<tr>
<td>increase both the firm’s and society’s costs</td>
<td>institutional settings it would be viewed as ungracious, if not positively immoral, for those</td>
</tr>
<tr>
<td>(perks and sloth), others will appear as mere</td>
<td>favored by regulation not to share some of the rents with those who granted them. This</td>
</tr>
<tr>
<td>transfers from the owner to stakeholders in the</td>
<td>can be viewed merely as one of the sources of regulatory capture. However, civil servants are</td>
</tr>
<tr>
<td>firm (Krueger 1974). A higher wage for workers is</td>
<td>not unaware that some government jobs are “wet” (with high potential for receiving such gifts),</td>
</tr>
<tr>
<td>the most obvious example. However, the literature</td>
<td>while others are “dry.” They will naturally spend resources to gain access to a “wet”</td>
</tr>
<tr>
<td>on rent-seeking reminds us that where there are</td>
<td>position, and—as in the first column—real social costs are incurred.</td>
</tr>
<tr>
<td>rents to be had, people will expend real resources</td>
<td></td>
</tr>
<tr>
<td>in their pursuit. One need not accept the extreme</td>
<td></td>
</tr>
<tr>
<td>version (that the resources expended will fully</td>
<td></td>
</tr>
<tr>
<td>absorb the rents) to recognize that much, if not all,</td>
<td></td>
</tr>
<tr>
<td>of the rectangle represents real welfare losses.</td>
<td></td>
</tr>
</tbody>
</table>

a. A similar story can be told for an owner-operator. As he becomes rich on monopoly profits, he will be wise to take some of the increased income in increased leisure and perquisites, with effects similar to those described in the text. Given the monomania of many big entrepreneurs, this may often be of minimal import for the first generation of tycoons but will take on increasing importance for their heirs.

b. Regulating a multiproduct firm (such as telecommunications) by imposing a single average price allows the firm to cross-subsidize, charging high prices in noncompetitive markets (such as home service) and precluding entry via lower prices in more competitive markets.
Figure 1. The Algebra of Little Triangles and Big Rectangles: Welfare Loss from the Presence of Market Power

The additional welfare loss, from higher costs (big rectangle) is indicated by shaded area \( B \). The welfare loss from higher prices (little triangle) is indicated by shaded area \( A \). The price elasticity of demand is defined as \( \epsilon = (\Delta Q/Q)/(\Delta P/P) \). The price distortion factor is defined as \( d = \Delta P/P \). Then \( \Delta Q = \epsilon dQ = Q^* - Q^M, \Delta P = dP, \) and \( Q^M = Q^* - \epsilon dQ \). The respective areas become:

\[
\begin{align*}
B &= \frac{1}{2} \Delta P \Delta Q \\
B &= \frac{1}{2} (dP^*)(\epsilon dQ^*) \\
B &= \frac{1}{2} \epsilon dP^*Q^* \\
\end{align*}
\]

\[
\begin{align*}
A + B &= \Delta P Q^M + \frac{1}{2} \Delta P \Delta Q \\
A + B &= (dP^*)(Q^* - \epsilon dQ^*) + \frac{1}{2} \epsilon d^2P^*Q^* \\
A + B &= dP^*Q^* - \frac{1}{2} \epsilon d^2P^*Q^* \\
\end{align*}
\]

(The triangle algebra is from Harberger; the rectangle algebra is added.) For example, let elasticity equal 1 and distortion equal 10 percent. Then \( B = 0.5 \) percent of revenues and \( A + B = 9.5 \) percent of revenues.

Equation 1 shows that the welfare loss is proportional to the total revenues of the enterprise, where the factor of proportionality is one-half the product of the
price elasticity of demand and the square of the price-distortion factor (that is, the ratio of the price increase to the base price). Thus if elasticity is in the vicinity of 1, the welfare cost of a monopolist’s raising prices by 10 percent is in the vicinity of half of 1 percent of revenues \((0.5 \cdot 1 \cdot 0.1 \cdot 0.1)\). This vicinity is clearly a low-rent district and remains so even if the estimate is off considerably. One might therefore wonder why anyone would undertake empirical work on the subject; it is not surprising to find that the original Harberger estimate of the cost of monopoly in the United States was 0.1 percent of gross national product (GNP) and that subsequent refinements have boosted that figure to only a few percent. (For a survey of the literature, see Scherer and Ross 1990.)

A gain of 1 percent of revenues may be small in relative terms, but it could nonetheless be large in absolute terms for, say, a large electric utility. How might the costs compare? That depends on the institutional mechanism chosen, but consider the classic U.S. model in which “cost plus fair rate of return” is determined in an advocacy setting. The first cost is that of the public utility commission; the second is the cost to the enterprise of preparing petitions, responses, and cases; and the third is that of the economists and lawyers hired as consultants. Although an institution that generates demand for economists cannot be all bad, two things seem clear: the opportunity costs are substantial, and they will be substantially higher in developing countries and formerly socialist countries, where the requisite talent is in short supply. Confining ourselves to the United States, I am not aware of any study that actually quantifies these costs, but would anyone be surprised if they were of the order of magnitude of 1 percent of revenues, or even higher? Costs might thus exceed benefits on this ground alone.

Matters look even bleaker if we take into account the second evil of regulation. This can take many forms, but consider only the classic American variant. If I reimburse you for costs, you have no reason to minimize any costs and every reason to maximize capital costs, since the absolute value of the “plus” will then rise (assuming that the “plus” rate of return is above the opportunity cost of investment and there is unexploited market power; Kahn 1992). Such cost- or X-inefficiencies give rise to big rectangles of welfare loss. A look at figure 1 or at Leibenstein’s (1976) Beyond Economic Man might convince one that the triangles are considerably smaller than the rectangles. Once again, a little algebra makes this clear. Equation 2 shows that the welfare loss from cost efficiency is, as a first approximation, proportional to revenues where the factor of proportionality is simply the cost-distortion factor (exactly analogous to the price-distortion factor above).\(^1\) The result is slightly more accurate if we subtract the second term in the equation, which is simply the little triangle. Accordingly, if costs rise by 10 percent, the welfare loss is 9.5 percent of revenues \((0.10 \text{ minus the corresponding triangle})\). If, therefore, a costless regulatory body succeeded in reducing the exercise of market power by 10 percent but at the same time precipitated a rise in costs of 10 percent, the net welfare loss from regulation would be about 9 percent of total sales.\(^2\) There is, of course, no reason why the price and cost distortion factors should be of equal magnitude, but the point is
that the cost distortion is far more important, and even a little cost inefficiency will wipe out the gains from a sharp price cut.

If these back-of-the envelope calculations are even roughly correct, it is easy to see how the costs of regulation could exceed the benefits. In fact, it is difficult to see how they could not exceed the benefits. Accordingly, it is not mere indifference that leads some policymakers to argue against regulation in the developing countries and the formerly socialist countries. We are most unlikely to make things better in any event, the reasoning goes, so why not simply relax and enjoy it—let the robber-barons make their piles now and generate the savings needed for growth? The markets will eventually catch up to them, and later generations will dissipate their wealth. The reader may correctly infer that I reject that conclusion. Nonetheless, these are some important lessons here, which I recapitulate below.

**The case of large distortions.** The first weakness in the no-regulation argument is that it is predicated on relatively modest price distortions. In a Harberger-type study of an entire industrial economy, this is reasonable, since many sectors have no distortions and problems in the potentially most distorted sectors have already been ameliorated by regulation. If, however, we want to know what the potential gains are in moving from zero to perfect regulation in a particular sector, the magnitude of the distortion may be considerably larger. This is obvious from figure 1: as the distorted price line \( PM \) rises, the triangle becomes larger and the rectangle smaller, so that at the extreme of zero production the rectangle disappears and the losses from allocative problems and X-inefficiency (corporate governance) problems become equal. No one would go to this extreme, and a natural limit is found at the distortion effected by a pure profit-maximizer. A model by Galal and others (forthcoming) reveals that the profit-maximizing price distortion is \( d/c^* \), where \( E \) is evaluated at \( Q^* \).

\[
(3) \quad d^* = \frac{1}{2e}.
\]

That is, for an elasticity of 1 (at the welfare maximum where price equals marginal cost), the profit maximizer would raise prices 50 percent, yielding (from equation 1) a welfare loss of 12.5 percent.

At this level of welfare loss, the elasticity component of the proportionality factor becomes important. There is an inverse relationship between the elasticity and the optimal price distortion: the more inelastic the demand (the smaller the elasticity), the more the profit maximizer will raise prices. Lower elasticity thus goes with higher distortions. The former reduces the welfare effect, but the latter increases it, and since the latter is squared, it dominates. Lower elasticities generate greater maximum potential allocative losses from monopoly (approaching 100 percent for elasticities in the teens), while higher elasticities generate lower losses (approaching 1 percent for high-single-digit elasticities). Accordingly, society has little to fear from monopolists facing elastic demand and much to fear from those facing inelastic demand.

186

*Appropriate Regulatory Technology*
How large might the relevant elasticities be? There is obviously considerable variance over sectors, countries, and estimation methods, but it is clear that in some critical industries they can be quite small. In sectors such as electric power, short-term demand is highly inelastic; values in the vicinity of 0.20 are not uncommon. In such a sector the optimal distortion is given by equation 3 to be 2.5, and the corresponding welfare loss (from equation 1) is 62.5 percent of revenues. This is certainly large.

Long-term elasticities, however, are always higher because people have time to adjust. A short-term elasticity of 0.2 might, for example, correspond to a long-term elasticity of 2.0, and the latter yields a welfare loss of only 6 percent. In such an industry, which measure of welfare loss is correct, 60 percent or 6 percent? The answer is a weighted average, where the weights are a function of the discount rate of the monopolist. The monopolist will maximize the net present value of a stream of profits, taking into account the fact that higher prices now yield lower demand later. Given the very high discount rate for entrepreneurial capital in developing countries (say, 30 percent), the first few years loom large in this decision, and the average can be closer to 60 than to 6 percent. In sum, little triangles are not always so little, and the scope for allocative gains from regulation is thus sometimes substantial.

In such sectors it is obviously quite possible to have substantial net gains from regulation so long as one avoids mechanisms that motivate cost inefficiencies or other distortions. In well-developed market economies such sectors will be largely confined to regulated public utilities, explaining low additional potential gains in Harbergerian-type studies. How will matters differ in developing countries and formerly socialist countries?

Industrial concentration in developing countries. Is the problem of market power more or less serious in developing countries than in industrial countries? To my knowledge, the empirical work does not allow a definitive answer to this question, but enough information is available to permit informed conjecture. Begin by decomposing the question into two parts. First, for a given industry operating in both developing and industrial countries, how does concentration vary? Second, what is the share of highly concentrated sectors in relation to the entire economy?

The first question can be further decomposed, since within an industry concentration depends tautologically on the relationship between plant size and market size. With other factors held constant, developing countries have smaller markets. If technology were standardized, all countries would have the same minimum efficient scale of plant, but larger and richer countries would have more plants and thus less concentration. Technology, however, does adjust, but not fully. The elasticity of plant size to market size has been shown across a wide range of studies to be approximately 0.5 (Caves 1989). Since plant size thus declines less rapidly than market size, empirical studies show rising concentration as we move to smaller economies (Scherer and Ross 1990, pp. 82–89).
But matters are more complicated than this. For example, the foregoing conclusion would predict less concentration in the largest developing countries than in the smallest industrial countries, but this need not be so because other things are not necessarily held constant. Sutton (1992) shows that although the inverse relationship between market size and concentration holds for industries with low expenditure on advertising, research and development, and other endogenous sunk costs, it does not hold where such costs are high. The same literature emphasizes that market structure also varies with the degree of price competition. Exogenous price competition (independent of market structure) depends on the ability of a small number of producers to collude. This in turn depends on the government's capacity to prevent such collusion and the ability of the press to publicize it. In many developing countries government policies (such as licensing requirements) have worked to increase, rather than to reduce, industrial concentration. That is, rather than prohibiting a firm from erecting barriers to entry, they erect them for it through licensing. Furthermore, enforcement mechanisms and other countervailing institutions are weak.

Other factors that affect concentration include openness to trade and the quality of the transport system. For any particular country a host of such factors enter in, but for present purposes it is only necessary to establish that most developing countries have small economies and that concentration will thus be larger for a given sector. Available empirical data are thin, but they support this view (Lee 1984).

The other half of the question may be answered more quickly. In developing countries agriculture accounts for a much larger share of gross domestic product (GDP) and manufacturing for a smaller share. Therefore the share of concentrated sectors will be smaller because of systemic differences in industrial structure.

In sum, the scope for welfare improvement will be larger because the modern sector is more concentrated, but that sector will generate a smaller fraction of GDP. How do these two offsetting tendencies net out? For the sake of argument, assume that the share of concentrated sectors is 10 percent of GDP. This is the order of magnitude of public enterprise sectors in developing countries, and these sectors overlap substantially with concentrated sectors (Jones and Mason 1982). Assume further that half of this is in utilities with short-term elasticities of 0.2 and long-term elasticities of 2.0, yielding welfare gains in the vicinity of 30 percent. Assume that the other half of the concentrated sector is in oligopolistic industries that imperfectly exploit their market power but have similar elasticities. The welfare gain could be something like 10 percent in the oligopolistic subsector, and the average for the sector would be 20 percent of total sales. Since concentrated sectors tend to be intermediate-intensive, 10 percent of value added would translate into sales of about 25 percent of GDP. The potential welfare gain would be 20 percent of that, or 5 percent of GDP. This is on the order of double what improved Harbergerian analysis yields for the United States and appears plausible (although I suspect it errs on the low side). Pending research, let us take it as a working stylized fact.
The next question is whether 5 percent of GDP is enough to worry about. Five percent of GDP is, after all, of the same order of magnitude as total corporate and personal income tax collections in the average developing country; it amounts to 28 percent of total taxes and thus—allowing for borrowing and deficit finance—to perhaps a fifth of total government expenditure. Big is in the eye of the beholder, but I see this as big.

**Industrial concentration in formerly socialist countries.** Data on industrial concentration in the formerly socialist countries have not been analyzed as thoroughly as for industrial countries, but available data from Russia permit the following preliminary generalizations.

- Enterprises are typically bigger than in the West, but markets are smaller, and individual *enterprises* are therefore more likely to dominate particular industries (IMF and others 1991, pp. 16–17, 36–40).
- Even worse, enterprises are widely incorporated into *groups* (for example, cement, fertilizer, or heavy equipment) so that at this level concentration frequently approaches 100 percent.
- Because transport networks are less developed, regional monopolies are common.
- Even where there are many producers, there is often only one middleman because the wholesale and retail trade sectors are monopolized.
- Competition from foreign trade is often minimal,
- Countervailing forces are weak or absent; consumer advocacy groups are nascent, antimonopoly and regulatory bodies are in their infancy, and the legal system is underdeveloped.

If we start with the fact that enterprise monopolies account for 30 to 40 percent of mining and manufacturing output, a crude adjustment for group monopoly as well as oligopoly suggests that highly concentrated industries must account for at least half, and probably two-thirds, of total output. Like the developing countries, formerly socialist countries have much more concentrated markets than is common in industrial countries. However, the share of modern sectors in GDP is much closer to that in industrial countries. Accordingly, it seems safe to conclude that imperfectly competitive markets are a considerably larger problem in formerly socialist countries than in developing countries and that the potential welfare gains discussed in the preceding section should be increased substantially in formerly socialist countries.

**The second evil of monopoly.** This far I have argued that little triangles might sometimes be large enough to worry about in both developing and formerly socialist countries. Now let us consider the rectangles. The second problem with the no-regulation argument is that it ignores the market failures of monopoly-induced sloth and rent-seeking. The resulting cost inefficiencies are measured in terms of big rectangles, so the cost of even modest distortions can be substantial.
How big might these rectangles be? Under the same conditions that yielded a 5 percent triangle in developing countries, and on the basis of equation 2, the rectangle would be about 15 percent of GDP. This calculation assumes that rents would be fully converted into costs, but 5 to 10 percent strikes me as a plausible estimate of the gains from eliminating market power in developing countries. In formerly socialist countries these gains would be considerably larger.

Once again, there are potential gains from well-designed regulatory mechanisms. As I discuss below, the ability of price regulation to deal with this problem is probably quite limited, but cost inefficiencies do provide an argument for some of the structural remedies in oligopolistic markets.

**The Politics of Regulation.** A final, and rather different, problem with the no-regulation argument is that it ignores the political reality of popular pressures to regulate, usually associated with calls for distributional equity. Such appeals are easily overstated in the context of developing countries. For example, the truly poor do not have access to electricity at all, or they have a single bare bulb dangling from the ceiling. The comparatively well-off have refrigerators and air conditioners and consume other energy-intensive products. Thus raising utility prices in such countries harms mostly the top third of the income distribution. If those increases are captured by the government (through taxes or public enterprise profits) and are used to help the poor, income distribution will improve. If, however, profits accrue to rich private owners, the transfer is still from middle-class consumers.

Helping the middle class is, of course, an important political goal; they are the ones who vote and who riot in the streets. Because the room for Pareto-efficient reallocations (which make someone better off without making someone else worse off) to this class is small, politicians naturally turn to what I (in Jones 1985) call politically efficient reallocations, which make someone better off without making anyone else aware that they are worse off. In this latter sense, Pareto-inefficient price regulation is politically efficient because the beneficiaries of lower prices are aware of it but the losers are not. Most governments are therefore likely to regulate, for better or worse.5

**Summary.** The no-regulation argument is something of a rhetorical straw man, set up to make the following points:

- The benefits of regulation are easily overstated because little triangles are often just that—little.
- The costs of regulation are often unrecognized and include both significant resource costs and potentially much larger regulation-induced distortions.
- The net benefits of regulation can easily be negative. The doctor’s creed, “first of all, do no harm,” should be adopted by regulators.
- There is, nonetheless, considerable scope for welfare-enhancing regulation, but the targets must be carefully selected, and mechanisms must be designed to minimize the costs in relation to the benefits.
Particular attention should be paid to the big rectangles associated with cost-inefficiencies, where most of the really big numbers lie. This applies to both targeting the benefits from removing market failures and avoiding the costs associated with government failures.

Price Regulation

Regulatory technology is characterized by diminishing marginal benefits and increasing marginal costs. The benefits from removing allocative inefficiencies (although not X-inefficiencies) diminish because of the squared distortion term. Eliminating the first 1 percent of a 20 percent allocative distortion is approximately forty times as beneficial as removing the last 1 percent (assuming unit elasticity). Costs rise because of information problems. It is not difficult to determine whether a company's prices are above the efficient level: international comparators convey a great deal of information. Figuring out, however, that a firm is a few percent above the efficient level is probably impossible. The trite but true lesson is to keep it simple: try to control the more egregious problems and leave the smaller issues aside. How do various regulatory mechanisms measure up under this criterion? I start with price regulation.

The Effect of Cost-Plus Pricing on Incentives

It should be apparent that U.S.-style cost-plus regulation does not fare very well. For one thing, it is conducted in an advocacy setting and incurs high costs to obtain the last few percentage points of accuracy. Even if this is cost-effective in the United States, it is highly unlikely to be of use in countries in which professional skills are scarce.

Leaving this problem aside, the fundamental flaw is that cost-plus pricing encourages cost inefficiency. Many incentive mechanisms have been proposed to deal with this problem by leaving some fraction of cost savings to the enterprise. (In this context, "incentive" refers to the problem of motivating the firm to act like a true profit maximizer in minimizing costs.) Laffont and Tirole (1993) point out that such schemes can be arrayed on a scale from zero to one according to the percentage of cost savings that accrue to the enterprise. At one extreme (see figure 2) are ex-post cost-plus contracts that give the firm no incentive to economize. At the other extreme, under a fixed price cap, the firm has full incentive to economize. Incentive schemes are anything in between. Those on the cost-plus end of the spectrum are termed low powered in their ability to induce efficiency, while those at the price-cap end are termed high powered. Note that the spectrum reflects a classic efficiency-equity tradeoff because all the benefits have to be given to the enterprise in order to maximize the quantum of benefits to be distributed.

In practice, the U.S. version of cost-plus pricing does not fall at the zero end of the spectrum. It is not ex post in the sense that the firm simply sends in bills and
is reimbursed for its costs (including the opportunity cost of capital). Rather, it is ex ante; the last period’s costs serve as the basis for the current period’s prices. This method generates some surplus, which in turn is used as the basis for determining the next period’s prices, and so forth. The resulting regulatory lag adds power to the scheme. If prices are fixed for a period of years, any cost savings in the interim accrue to the firm. If prices were fixed forever, it would be equivalent to a price cap and fully high powered. However, with a finite regulatory lag, the enterprise knows that when the next price is set, lower costs this period will be taken into account in setting a lower price next period. Some cost reductions will be appropriated by the government for consumers (or for itself in higher taxes), and the incentive effect is thus less than complete. Nonetheless, the longer the lag, the higher the power of the system.

Are Price Caps a Better Formula?

It is easy to see why the incentive power of regulation might be raised by switching to a price-cap formula such as the United Kingdom’s “rate of price inflation minus \( X \)” (RPI – \( X \)). If an enterprise is told that for the next, say, three years it can raise prices at no more than, say, RPI – 2, will it not have every incentive to lower costs? Costs, after all, are not even mentioned. The answer is
that this will be fully high powered only if that rate is reasonably expected to last forever. If not, then we must ask how $X$ will be determined for the next period. The manager might think that if profits were 30 percent, $X$ would rise significantly, as opposed to what it would be if profits were 10 percent. Accordingly, there is the same danger of a ratchet effect as with cost-plus pricing: the fact that profits may be expropriated in the future reduces the incentive to cut costs now. That costs are not explicit in the formula does not mean that they are not implicit in the $X$. If $X$ is not determined as a function of revealed profits and costs, how is it determined?

One advantage of $RPI - x$ is that it is typically associated with a relatively long lag. However, this dimension of choice can easily be added to cost-plus schemes. The proper comparison is between two managers in the same environment; one can raise prices 2 percent a year under cost-plus and the other can raise prices the same amount under $RPI - x$. I have yet to see a convincing argument as to why the two managers would behave any differently, given a common set of regulators and a common expectation of what figures regulators will examine.

A related advantage of $RPI - x$ is that it allows a longer lag because it explicitly adjusts for one of the major exogenous environmental variables (Vogelsang 1989). Allowing the price to vary with the rate of inflation thus promotes a longer lag before prices have to be adjusted. Given higher rates of inflation in developing and formerly socialist countries, this is no small advantage. Much the same thing, however, can be accomplished with a crude tool (such as the fuel-adjustment clause in U.S. regulations, which calls for a surcharge on electricity bills if the prices of energy inputs rise more than a specified amount). The adjustment can also be accomplished in a far more sophisticated way. In Chile the telephone company adjusts charges every two months on the basis of a Divisia index of the prices of inputs into each service (Galal 1992).

In sum, the incentive power of a price regulation system does not depend fundamentally on whether it is couched in cost-plus or $RPI - x$ terms. Rather, it depends on the length of the regulatory lag and the expectations of how prices will be adjusted at the end of that lag.

**Price Adjustment Techniques: New Zealand versus Chile**

One major advantage of $RPI - x$ is that in practice—although by no means necessarily—it has been associated with relatively simple implementation mechanisms. Perhaps the ultimate in simplicity is found in New Zealand, whose system has been described as "regulation without regulators." That is, there are no industry-specific regulatory bodies, and restraints are embodied in regulations. This authority seems to be used with considerable restraint. For example, in the case of telecommunications, the regulations provide that "the standard residential rental for a phone line will not rise faster than movements in the Consumer Price Index unless the profits of Telecom's regional operating companies are unreasonably impaired" (New Zealand, Ministry of Commerce,
It also stipulates that rural rates will not be higher than residential rates and includes a service requirement. Prices of other services are regulated only by competition or the threat thereof. The latter provision is particularly important and will be elaborated on in the next section.

The RPI - 0 pricing mechanism has some clear advantages. First, it is extremely economical in terms of the cost of regulation. It is not quite costless because there is a communications division within the Ministry of Commerce. However, this division is small, and the bulk of its time is spent on entry issues rather than pricing. Second, its cost-incentive power is quite high because of the indefinitely long regulatory lag with a fixed price in relation to inflation. Furthermore, the regulation gives the firm an escape clause that allows prices to rise faster if profits are "unreasonably impaired." Given the recent rapid rate of technological progress in the telecommunications industry and the scope for internal efficiencies in the recently privatized enterprise, this provision is unlikely to be needed for some time, but it presumably is a comfort to the firm to have it there just in case.

The scheme nonetheless falls short of being completely high powered. There is always the danger that some future government will impose price controls. A judicious manager will take this into account and will keep profits from rising to unreasonable levels. Ideally, of course, he would do this by lowering prices. Alternatively, he might not work quite so hard to exploit cost-cutting opportunities. Given the latter possibility, the scheme is something less than fully high powered. Furthermore, it does not attack the second evil of monopoly, instead leaving that corporate governance problem to the private sector.

Still, the New Zealand compromise has much to recommend it. The primary emphasis is on cost-efficiency incentives, with considerably weaker controls on the allocative inefficiencies of monopoly pricing. Given the relative sizes of rectangles and triangles and the importance of correcting the worst of the triangle abuses, the direction of the bias in the compromise cannot be questioned, although one might quibble about the magnitude.

Chile employs a rather different model (Galal 1992, pp. 11-13). The system can be described as cost-plus-fair-return because firms are allowed a rate of return equal to the risk-free rate plus a premium based on the systematic risk of the industry and the difference between the risk-free rate and the return on a diversified investment portfolio. (It could as well be described as RPI - x because the price cap is adjusted every two months to reflect inflation.) Which phrase is chosen is immaterial because the real distinction lies elsewhere.

The first critical feature is specificity. If New Zealand is near the zero end of the vagueness-specificity continuum, Chile is at the opposite end. For one thing, the adjustment period is explicit and reflects a long lag of five years. Moreover, a range of regulatory technology is spelled out in law. In addition to the sophisticated fair-rate-of-return and inflation-adjustment mechanisms already described, long-run marginal costs are calculated in the context of a five-year investment plan (designed to minimize the system costs of meeting projected demand), markups from marginal to average costs are apportioned via Ramsey...
pricing, and so on. There is more, but the general point of sophisticated specificity is clear. Note that this mechanism reduces the firm's uncertainty and focuses any debate on technological parameters. This approach is similar to that of the U.S. model, which allows for only "reasonable" costs. The difference lies in the ex-ante specificity about how the parameters are to be determined. In the United States it is in an advocacy proceeding guided by general principles, but in Chile the principles are detailed in advance.

The second critical point is that the model relies on data not simply from the firms themselves but from similar firms in other countries. In principle, such benchmark or yardstick pricing is unassailable. Given a sufficiently large set of sufficiently similar firms, one could readily use econometric techniques to establish a best-practice technological frontier. Add data on local prices and the cost of capital and it would be possible to generate the right price for the company. The problem, of course, is that there is a shortage of sufficiently similar firms, and adjustment must be made for local conditions—that is, using data from the regulated firms themselves. In the case of electricity distribution in Chile, this problem is minimized because there is a number of similar distribution networks. Data are gathered from the set as a whole, and a local best-practice benchmark is established (Galal and others forthcoming). Forced to meet prices generated by this frontier, a given firm has no incentive to shirk so long as its own costs are a small fraction of the total.

An appealing property of this approach is that, unlike any of the other schemes discussed, it solves the corporate governance problem so long as at least one firm in the sample has solved it. For Chile Telecom this condition does not hold because the comparators are imperfect (foreign) and are drawn from a set of regulated private firms or public enterprises, where cost minimization is not likely. Nonetheless, the technological benchmark model is clearly a major step in the right direction.

A third feature of the system is that if the regulator and the firm disagree, they can appeal to a technical arbitration board. Finally, the system is reasonably economical as to the cost of regulation; the bulk of the work consists of a study that is produced every five years.

The New Zealand and Chilean systems are quite different, but they define the boundaries of what might be practical. The system used in New Zealand is extremely simple. It minimizes the costs of both of the major regulatory failures, at the expense of the two major market failures. Chile's system makes considerable progress on the two market failures as well, but at the expense of some complexity and information requirements. Many intermediate solutions are possible. For example, New Zealand's system could be designed so that $X$ is estimated on the basis of exogenous data on the rate of productivity change in the industry and adjusted for changes in relative price trends. I would suggest that the feasible choice for regulatory design lies on the continuum between the two countries. The precise point on that continuum would be a function of specific country and industry conditions.
Underinvestment: Expropriation versus Exploitation

Thus far I have focused on the way in which regulatory expectations can lead firms to raise costs. However, they can also induce underinvestment if it is feared that regulators will expropriate future earnings. In a monopoly environment, underinvestment is a serious economic problem. A World Bank study (Galal and others forthcoming) of divestiture in Chile, Malaysia, Mexico, and the United Kingdom found that much of the welfare gain came from an expansion of investment and that most of the benefits accrued to consumers. Levy and Spiller (forthcoming) study regulation in five countries, with a primary focus on observed underinvestment.

There are, of course, other explanations for underinvestment than the fear that regulators will expropriate future earnings. First, the company may have underestimated demand. Second, public enterprises underinvest when they are bound by public borrowing constraints rather than by the economic merits of their particular projects. Finally, underinvestment is exactly what a profit-maximizing monopolist would do to reduce quantity, create excess demand, and increase prices under a friendly regulatory regime. In this guise, underinvestment is simply one manifestation of the first evil of monopoly.

Whatever the source, the associated welfare loss is measured by the relevant little triangle. The operations of France Telecom in the 1960s provide a nice example of this phenomenon (Anastasopolus 1973, p. 164). Despite offering abysmal service, the company was consistently highly profitable. The actual rents, which were high, were absorbed by higher costs, thus illustrating the second evil of monopoly.

In other countries the same phenomenon also illustrates the third (rent-seeking) evil of monopoly. Excess demand plus price regulation means that some people are willing to pay a considerable amount for a phone line. In some instances a legal market is created to reallocate old lines and auction off new lines. In others the problem is resolved by queuing and by bringing pressure. In many developing countries, however, it is necessary to pay several hundred or several thousand dollars to someone with the right connections. These rents may go to employees or to political appointees, but they do not go to finance the expansion of telephone service. In such a setting an expansion of service would damage prominent decisionmakers and—even if they are not strictly profit-maximizers—this likelihood may explain why the shortage exists and persists. Whatever the reason, underinvestment is expensive. What can be done about it?

Mechanisms to Create Positive Regulatory Expectations

In cases where underinvestment reflects a fear of expropriation, Levy and Spiller (in this volume) emphasize the importance of institutional factors:

For a regulatory system to encourage private investment, three complementary mechanisms must be available to restrain arbitrary adminis-
trative action: substantive restraints on the discretion of the regulator, embedded in the design of the regulatory system; formal or informal procedural constraints on changing the regulatory system; and institutions that enforce these substantive and procedural constraints.

These conditions are profoundly important but are covered in other papers in this volume.

Another mechanism is reputational. It has been suggested that the threat of expropriation can be mitigated by covenants—or conditions—in lending agreements with international financial institutions. Those who doubt the potential efficacy of this mechanism must explain why else the countries involved in the Latin American debt crisis did not default on their loans. The answer presumably has to do with reputation, which is a precious international commodity. Still, one might wonder at what point international agencies would go to the mat to enforce a pricing contract. Even if this prevented a wholesale expropriation, at what margin would the deterrent operate? Given the inevitable accounting and legal questions, would it be possible to deter, say 25 percent exploitation?

The strength of reputational deterrents will vary with whose property is being expropriated. By granting part or all of the license to foreigners or selling shares to dispersed private domestic shareholders, the probability of expropriation is reduced. Both mechanisms were used in the U.K. divestiture.

If one were not sure which problem would crop up, or simply wanted a little insurance, a more complete contract should be considered. This statement rests on the principle that achieving two goals requires two instruments. A contract will be required that explicitly incorporates pricing and investment rules and that makes expansion an explicit condition of the license. In the United States a utility is typically expected to provide service to anyone willing to pay. This is obviously not feasible in developing countries, but Mexico applies the same principle by requiring that the number of lines in service must expand by 12 percent a year, that all towns with more than 500 people must have service by a fixed date, that the number of public telephones must increase at a specified rate, and that the maximum waiting time for a phone must be reduced to six months by 1995 and to one month by 2000 (Tandon forthcoming).

This seems a sensible solution. It is particularly appealing because once the investment is in place, the potential for enhancing profits by cutting quantity is substantially reduced. One drawback is that the firm would require a higher expected return in this case, and if credibility were an issue, that return might be very high indeed.

Laffont and Tirole (1993, pp. 100–03) suggest four other mechanisms, three of which are straightforward: reimburse a high fraction of the investment costs; rely on repetitive regulatory interaction to develop credibility over time; and, where possible, allow the firm to use its capital to produce nonregulated commodities. A fourth—somewhat facetious—suggestion would entail installing a captured regulator. The same thing can be accomplished by granting the license...
to a relative or compatriot of the president. The problem may then be overin-
vestment (if AJW conditions obtain).

Although this course is not to be recommended, it does serve to remind us that
capture will be the likely outcome in many countries. In fact, I would hazard a
guess that the phenomenon of capture and resulting exploitation will be the
problem in considerably more countries than will be the danger of expropria-
tion. This may simply be the result of bias in the sample of countries I know
well. However, there is some logic in the inherent imbalance in power relation-
ships between regulator and regulated. Large public utilities (and other firms
with sufficient market power to worry about) are going to be owned or run—or
both—by rich, powerful individuals. Regulatory commissions are going to be
run by technocrats or midlevel politicians. An amusing illustration occurred in
Chile under Allende, when the two phone companies were run by generals and
the regulatory agency by a colonel.

Although no wholly satisfactory solution to underinvestment is available,
some combination of the foregoing ideas may substantially ameliorate the prob-
lem in particular contexts.

Other Regulatory Issues

In part because of space limitations and in part because they are obvious, regula-
tory issues other than natural monopolies will be dealt with more perfunctorily.
They are included only to provide a balanced perspective. Others, especially
problems of antitrust and aggregate concentration regulations, however, are
quite complex and are decidedly underresearched in the context of developing
and formerly socialist countries. They are included to suggest further areas of
research.

Structural Remedies

Structural remedies alter market structures in ways that promote or simulate
competition. They are by definition institutionally appropriate, since they pri-
marily require one-time policy actions rather than ongoing monitoring.

Optimal regulatory policy. This is one area that needs no research; the
simplest and most effective regulatory mechanism is competition.

Tariff policy. A considerable share of the concentrated industries in develop-
and formerly socialist countries produces tradables. In this case the simplest
solution is to reduce tariff protection. I do not suggest ignoring the importance
of infant industry regulation. As we know from the experiences of the Republic
of Korea and Japan, such protection can be quite effective if it is done right.
Doing it right means not extending the protection to teenage—let alone
geriatric—industries. Provide a firm schedule of tariff reductions tailored to the conditions of individual industries, and tell the babies they have a fixed number of years to grow up.

Reducing government barriers to entry. Domestic competition can also be enhanced in ways that are not only administratively economical but actually save resources. Most developing and formerly socialist governments discourage rather than promote competition by imposing such entry barriers as general licensing requirements and cumbersome import procedures. They also prevent competitors from impinging on the territory of natural monopolies, which have all too often been de facto unnatural monopolies (such as provision of telephone equipment or cellular services). Such barriers should be eliminated, not on a fixed schedule, but immediately. The problem is not complex in economic terms, but it does face considerable political barriers.

Breaking up large units in the formerly socialist countries. Competition can also be fostered without new entrepreneurship by breaking up large existing enterprises. The potential here is particularly large in the formerly socialist countries, where, as noted earlier, economic ideology promoted huge units. This step will require careful attention to economies of scale and scope. The danger is apparent from the example of Mongolia cited earlier. Given the country’s human resource endowment, the fact that competition will not be materially affected, and the very limited scale of telephone penetration in the country, it is highly likely that welfare will be reduced rather than enhanced by this policy. Substituting blind capitalist ideology for blind socialist ideology is not necessarily an improvement. The intermediate step of breaking the company into semiautonomous profit centers could both improve current corporate governance and help prepare for eventual dismantling without sacrificing economies of scale and scope.

Mongolia, however, is almost certainly the exception rather than the rule. The formerly socialist countries are littered with conglomerate behemoths that cover huge areas and produce everything from soup to nuts. In such cases a few fences and a little legal work could yield clear efficiency gains, unless economies of scope are considerably larger in the formerly socialist countries than in the United States.

Much more important than breaking up companies is the need to break up groups of companies. In Russia today a critical battle is being fought to prevent the reconstitution of the old branch ministries as holding companies that control shares in subordinate enterprises. If this battle is lost, the privatization process in the short run will have changed nothing internal to the public enterprise sector: the same old people will be making the same old decisions in the same old way. To be sure, they will eventually have to respond to market forces, but that could take a long time.
BREAKING UP LARGE UNITS IN DEVELOPING COUNTRIES. The scope for breakup in developing countries is less, but it is by no means absent. Egypt's large public enterprise sector has historically been organized into vertically oriented groups. For example, all fertilizer companies were in one holding company and all cement companies in another. The government has decided to commercialize most of these enterprises as a first step toward privatization. However, giving the firms a profit orientation while grouping them together under one controlling body has obvious dangers for consumers and less obvious dangers of perpetuating old management with old people under a new name. As a partial solution, the government is committed to reducing the number of holding companies substantially and reassigning portfolios so that, for example, cement and fertilizer firms are spread across holding companies.

In addition, the so-called natural monopolies appear less natural with the progress of technology. In the case of electricity the core grid is necessarily a monopoly, but the separate generating plants are not. For example, in Chile, which followed this model, the generating units were free to compete in selling power to the grid. The separate distribution units, although noncompetitive, at least gave regulators a comparative basis for regulating prices and costs.

Threat of Entry

The literature on contestable markets has taught us to look not only at actual but also at potential market structure. Market power may not be fully exploited if there is fear that high prices and high profits will attract entry. In the regulatory context, this line of reasoning is extended to include the threat of regulation. Thus it is argued that New Zealand's otherwise rather toothless price regulation scheme nonetheless has a significant constraining effect on the exercise of market power. According to an official of the Electricity Corporation:

Unless we can be confident that what we are doing will stand up to robust scrutiny and price on that basis, then what we are doing is just extracting monopoly rents. We'd be proving everyone right who says we are just a dirty monopoly and we'll end up price regulated. That is the last thing we want. (quoted in Spicer and others 1992, p. 145)

Similar claims have been made to explain the pricing of local telecommunications services. Determining how much of this is self-serving rhetoric and how much a reflection of actual pricing policies is a subject for analysis. Pending that, only note that threat of entry or more rigorous regulation can have an attenuating effect on otherwise unregulated pricing behavior.

Aggregate Concentration

This paper has emphasized concentration in particular markets. However, there is another important form of concentration: the share of business groups in the
total economy. Well-known examples include Japan’s prewar *zaibatsu* and post-war *keiretsu*, Korea’s *jaebul*, India’s industrial houses, and Pakistan’s twenty-three families. Such aggregate concentration has been of concern in the United States at least since the time of Thomas Jefferson. Berle and Means (1932) calculated that the 100 largest nonfinancial corporations held 49 percent of nonfinancial corporate assets and that if their relative growth rate continued, they would hold 100 percent by 1972. As with most such Malthusian arguments, this one failed to materialize (Scherer and Ross 1990). It has proved exceedingly difficult to identify any deleterious conduct resulting from this particular structure. That is, if aggregate concentration cannot be found to result in the exercise of market power, why worry about it? This is why one observer has termed aggregate concentration “a phenomenon in search of significance” (Brozen 1982).

In developing and formerly socialist countries, however, matters may be considerably different. If a particular family or business group is large in relation to the entire economy, it can use its power to gain privileged access to scarce resources. Such resources include underpriced credit, import licenses, tariff protection, and a host of favors granted by the government and the private sector in disequilibrium markets (Leff 1978, p. 666). As Leff notes, “The group pattern of industrial organization is readily understood as a microeconomic response to well-known conditions of market failure in the less developed countries. . . . The group can be conceptualized as an organizational structure for appropriating quasi-rents which accrue from access to scarce and imperfectly marketed inputs.”

 Aggregate concentration may not be a problem in industrial countries in which markets are closer to equilibrium, but it can be serious in countries in which markets are distorted. The locus of the problem is factor markets, where nonprice rationing of credit and other government-controlled inputs conveys unfair advantage across many product markets, even where those product markets themselves are uncontrolled. This issue deserves considerably more attention than it has received thus far.

**Conclusions**

An examination of the issues involved in adapting regulatory technology to developing and formerly socialist countries allows us to draw the following conclusions.

- Markets are more distorted in developing countries and formerly socialist countries because they are generally smaller in relation to minimum efficient plant size; because they have made less progress in developing institutions to deal with market power; because the room for rent-seeking behavior is greater; and because information is less widely available. If a Harbergerian framework yields estimates of the allocative inefficiency losses attributable
to market power of a few percent in the United States, it might amount to 5 percent in developing countries and considerably more in formerly socialist countries.

- The losses attributable to X-inefficiencies or corporate governance failures are considerably larger, perhaps on the order of 10 percent of GDP in developing countries and, again, considerably higher in formerly socialist countries.

These estimates, based on simple linear models and highly generalized stylized facts, are imperfect. Nonetheless, they suggest that without much more work, one cannot simply dismiss out of hand the need for regulation in these countries on the basis of the finding that the gains from further improvements are small in the United States.

Turning to the cost side,

- The costs of intervention are also higher in developing and formerly socialist countries. The opportunity cost of scarce human capital devoted to regulation is higher, and the potential for adding new distortions is greater. Regulatory design should therefore focus on minimizing these costs. In particular, since allocative costs are an exponential function of the magnitude of the price distortion, the marginal benefits of regulation decline with more regulation, and the costs rise because of increasing information requirements. It is therefore vital to keep regulatory mechanisms simple and to focus on controlling the most serious abuses.

The implications of this philosophy suggest that:

- The U.S. version of cost-plus regulation is not recommended because the direct costs of regulation are high in an advocacy setting (the first government failure) and because cost incentives are low powered (the second government failure).
- The New Zealand Rpi - 0 system neatly solves both of the government failure problems but may have only marginal impact on the allocative and corporate governance forms of market failure.
- Chile has resolved both market failure problems and the second government failure problem, but at increased cost (the first government failure).
- New Zealand's effort represents the minimum that should be done; Chile's, the maximum that should be attempted. Most countries will fall somewhere in between, with a relatively long regulatory lag and X set using as much exogenous data as possible.

Other regulatory problems, including competition and aggregate concentration, are dealt with more briefly.

- Competition is the most effective way to regulate market power. Structural means of ensuring a competitive marketplace include tariff policy, entry policy, and the breakup of large units.
• Given the institutional endowment of developing and formerly socialist countries, antitrust legislation is likely to be extremely costly in terms of market failure. The income elasticity of demand for such legislation is accordingly very high, and it can be argued that there should be little effective demand below, say, $2,000 per capita.

• Aggregate concentration conveys market power across product markets through privileged access to scarce credit and government licenses. This problem deserves considerably more attention.

In sum, institutional and factor endowments in developing and formerly socialist countries differ from those in industrial countries. This means that the nature of the problem changes. For example, both aggregate and market concentration are more serious problems in developing countries and formerly socialist countries. It also means that the tools available for solving the problems are scarce and must be used sparingly.

Notes

1. The distortion factor is not squared, which explains why triangles are small and rectangles are large in the case of small distortions. For large distortions, see equation 3 and the accompanying explanation.

2. In figure 1, add another horizontal line 10 percent above PM and let that be the unregulated price; let PM be the price and cost achieved by regulation, and let P* be the competitive price and cost. The loss would be exactly 9 percent.

3. India’s GNP is about ten times that of Israel, and recent work suggests that China may have the world’s third largest economy.

4. A sample of eighty-two developing countries shows that total tax collections average 18.1 percent of GDP, with about a third each from income, foreign trade, and other domestic taxes (Burgess and Stern 1992, pp. 14–17).

5. Bradburd (1992) provides schedules of the losses associated with different assumptions and offers significant refinements. Bradburd’s bottom line is somewhat closer to the pure no-regulation view. This is attributable in part to a difference in subjective judgment as to just what constitutes “big” in a policy context, in part to my placing somewhat more emphasis on short-term elasticities, and in part to my giving more weight to the second evil of monopoly.

6. The term was used by Guy Beatson in a speech at the World Bank in December 1992.

7. The 1986 Commerce Act “makes provision for the imposition of price control generally, or on particular firms, or even specific products and services, in circumstances where the Minister of Commerce is satisfied that conditions of effective competition do not exist and control is necessary; to protect users or consumers or, as the case may be, suppliers” (McCabe forthcoming).

References


Jones


McCabe, Patrick G. Forthcoming. "New Zealand: The Unique Experiment in Deregulation."


COMMENT ON "APPROPRIATE REGULATORY TECHNOLOGY," BY JONES

Henry Ergas

The past decade has been an age of regulatory reform. Far-reaching change has occurred in perceptions of the legitimate goals and instruments of public regulation. The shift in perception has been accompanied by new approaches to the design and functioning of regulatory institutions. Professor Jones identifies these approaches and draws lessons for policymakers, and there is little in his paper I disagree with. Yet it is difficult to avoid the impression that the paper is written at a level of abstraction quite far removed from the concrete problems regulation inevitably involves. It is on some of these operational issues that I focus my comments.

Let me start by echoing the sensible point that economic regulation should concentrate on the areas in which the scope for the abuse of market power is greatest. It is intuitive that these areas will have two characteristics: de facto nontradability and high barriers to discourage the emergence of new domestic suppliers. Because these characteristics are most marked in the activities conventionally regarded as public utilities, it is here that the case for economic regulation is strongest.1

It is possible to argue that even in these industries the costs of regulation are likely to outweigh the benefits, but Jones is quite right in regarding this argument as a straw man. It is not easy to conceive of an electrical or water utility subject only to those rights and obligations which apply to corporations that are not monopolies. In practical terms the question is not whether these industries are regulated but, rather, how the regulation works.

Jones considers this question in terms of the "technology" of regulation and the choices this technology offers. The analogy is a nice one, but one would have expected it to lead Jones along a somewhat different path. It is convenient to think of a technology as involving a relationship between outputs and inputs, so...
that identifying it requires (a) specification of the outputs and (b) assessment of
the impact on these outputs from both the "control variables" (those factors
under the regulators' direct control) and those variables not directly controlled
by the regulator, which form the context in which regulation operates.

The paper does not take this approach, and as a consequence, there is not as
full a discussion as there might be of the outputs that regulation is intended to
secure, of the "tool kit" available for this purpose, or of the extent to which the
context shapes the outcomes actually achieved. Our understanding of each of
these and of their implications for the design of regulatory institutions has, I
believe, changed significantly over the past decade. Two elements of this evolu-
tion are worth emphasizing.

There is, to begin with, a clearer appreciation of the impact that context has
on regulatory performance. Let me illustrate this by reference to my work on
Telecom Australia and its Swedish counterpart (Ergas, Ralph, and Sivakumar).
Both were, at the time, state-owned monopolies, both operated in geographi-
cally similar service territories, and both were largely self-regulating. Yet unit
costs were about 30 percent lower in Sweden, as were prices, while the range
and grade of service provided seemed very high. Undoubtedly, many factors
were at work, but not the least of these was Sweden's greater openness to
competition. International trade did not affect Swedish Telecom directly, but it
created an environment in which the largest users were acutely conscious of the
cost and quality of the service they received—a consciousness apparent in the
intense scrutiny that the Swedish parliament has traditionally reserved for public
monopolies.

By contrast, the high levels of protection that had long characterized the
Australian economy engendered a cost-plus mentality in which each firm toler-
ated the inefficiencies of others. It was only once protection had been signifi-
cantly decreased that concern about the costs imposed by the nontraded sectors
moved to the top of the political agenda, forcing Telecom into a far-reaching
effort at rationalization. Australia's experience in this respect is by no means
unique. It should be more widely recognized that opening economies to interna-
tional competition can have an impact that extends well beyond the sectors
directly affected.

But changing the context for regulation is not enough; reform must at the
same time extend to the scope of regulation and to the instruments and processes
on which it relies. Looking at the member countries of the Organization for
Economic Cooperation and Development (OECD) that have gone farthest in this
direction over the past decade, four common elements emerge. (For a useful
discussion of these, see Evans 1992.)

The first element is the separation of policy, regulatory, and operational
responsibilities. Fifteen years ago most public utilities outside North America
were effectively self-regulating (Foster 1992). The result was a murky accounta-
bility, as well as enormous (and frequently exploited) scope for abuse. Correct-
ing this situation has involved transferring regulatory powers from operating
entities to independent bodies while clarifying those matters which, because they involve political accountability, are best considered policy issues and thus within the proper remit of government.

Second, this separation of functions has been accompanied by an attempt to clarify the goals to be sought. A major criticism of the utilities, as of the nationalized industries more generally, was that the objectives they were intended to pursue were never properly set out. On the rare occasions when attempts were made to do so, the formulations did not lend themselves to the systematic assessment of performance. The trend now is to set specific goals for the regulator (in the enabling legislation) and for the regulated suppliers (both in the legislation and in operating licenses).

A third element is the recognition that one of the regulator’s goals is the promotion of competition. Just a decade ago, Littlechild (1983, p. 7) wrote that “competition is indisputably the most effective means—perhaps ultimately the only effective means—of protecting consumers against monopoly power. Regulation is essentially a means of preventing the worst excesses of monopoly; it is not a substitute for competition. It is a means of ‘holding the fort’ until competition arrives.” The central elements of this view—the emphasis on consumer interest as the primary concern and the reliance on competition to achieve it, even in industries traditionally regarded as largely involving natural monopolies—have received a far greater degree of acceptance than seemed likely at the time.

The fourth and last element is an emphasis on designing instruments that preserve the incentives the regulated entity has for efficient operation and that do not require complex monitoring procedures which are vulnerable to asymmetries of information and bargaining power. Practical examples include price-cap regulation, auctions of scarce resources (such as radio frequencies), reverse burden-of-proof obligations (on firms suspected of predatory pricing), and last-offer arbitration for interconnect pricing.2

In principle, all this should lead to a transition from regulation based on the power to make arbitrary decisions to regulation based on clear rights and obligations. Yet from the experience to date it is difficult to conclude that this is actually what we are getting. Rather, reformed regulation appears to be a complex and frequently burdensome process that is characterized by some tendency to expand rather than contract and by persistent concerns about the quality and predictability of the regulatory process. The United Kingdom is a case in point: regulation of electric power and gas is mired in conflict; water regulation has been controversial from the start; and the chief executive officer of British Telecom has said that the main effect of regulation has been to prevent the company from reducing its prices.

This is, to some extent, inevitable; after all, as Mao Zedong might have said, regulation is not a tea party. But it also reflects new strains that regulatory reform has built into the regulator’s task and the difficulties these strains have created for regulatory institutions. Three factors are especially important in this respect.
The first is that introducing competition into industries such as utilities, although undoubtedly worth doing, imposes a heavy cost in supplementary regulation. Consider the regulator's mission in terms of the following kinds of issues: industry structure, conduct, and performance.

- Issues of *structure* involve the delineation of boundaries between competitive, less-competitive, and noncompetitive services; the right to provide a particular service or the obligation not to provide others; and the definition of standards and arrangements for networking among different service providers.
- Issues of *conduct* involve preventing monopoly operations and thwarting unfair trading.
- Issues of *performance* center on prices, quality, and the range of services provided and also include compliance with obligations to provide common carriage, observe nondiscrimination, and provide uneconomic services.

In a monopoly setting, regulation largely involves the control of performance. The transition to competition expands the regulatory agenda into structure and conduct without substantially reducing the burden of monitoring what the dominant firm actually does. It is only once competition has become securely established and workably effective that the control of performance can cease, but this can take a long time and may never occur.

The second factor is that the transition to competition not only expands the regulatory agenda but also makes it far more complex. It may well be that, as Foster (1992, p. 167) has argued, “the most important innovation [in recent regulatory policy] has been the realization that there is no compelling reason why the [natural] monopolist should have the exclusive right to use its distribution network,” but the definition of how and on what terms third-party use of common carrier networks should occur has stretched (where it has not exceeded) the capacities of every regulatory institution. Equally complex issues have arisen in the control of discrimination and the prevention of conduct that is predatory or in other ways exclusionary. In each of these areas, the interests of the dominant player conflict sharply with the regulatory goal of allowing competition to develop, and the incentives to distort information and in other ways manipulate the regulatory process are even greater than they were in the past.

The final factor is the realization that the new instruments of regulation, although they have undoubtedly helped, have not proved as simple to implement as was frequently suggested at the outset. Price-cap regulation is a case in point; it raises complex issues about the level of the cap, the services to be covered, the treatment of rebates and discounts, the extent of cost flowthrough, the monitoring of service quality, and the treatment of demand and supply shocks. There is no doubt in my mind that it is far superior to regulation through direct control of costs or through capping rates of return, but it would be wrong to suggest that it is a “kiss and forget” solution.
In short, the regulator’s lot is not an easy one. The critical problem is that the issues which the new environment brings to the top of the regulatory agenda are extremely difficult, if not impossible, to deal with in the absence of a reasonable degree of regulatory discretion. For one thing, the “facts” involved in these decisions do not speak for themselves; they are often poorly defined and only exist as a result of complex judgments (for example, about the determination of costs). Moreover, the decisions typically involve tradeoffs between policy goals, so that assessment of these facts involves a significant measure of discretion. And finally, decisions are rarely permanent and will have to be reviewed as experience and technology evolve.

Discretion, however, creates vast problems. Economists generally think of the outputs of regulation in terms of the goals listed above, but regulatory legitimacy is also a goal. Decisions must be made that do not impose regulatory costs where they are not needed and that offer solutions that are timely, informed, and authoritative. It is important that the regulatory structure be predictable, consistent (internally and with other arms of government), and accountable for its decisions.

The greater the degree of discretion and the more difficult these criteria are to meet, the greater the risk that the regulator will pursue his own or interested parties’ agendas and that inconsistencies will develop between the rules affecting resource allocation in the regulated area and those at work in the economy as a whole. Minimizing these risks requires striking a middle course—one that I would locate somewhere between New Zealand’s approach and that adopted in the United Kingdom.

New Zealand spells out the obligations of the regulated entity in the operating license, which is largely protected from modification. Enforcement, as with any other contract, is vested in the courts. This approach assumes that complaints will be forthcoming when and where they ought to be and that the issues involved can be resolved in a court of law. Both assumptions are questionable. (On the limits to justiciability, see Sunstein 1990; Rosenberg 1990.) Indeed, the New Zealand courts have not proved at all capable of determining the issues that enforcement of the licenses has raised; a recent high court decision on the terms of interconnect seems to confirm that the formerly regulated monopolies are virtually free of effective control.

By contrast, the U.K approach gives industry-specific regulators a great deal of discretion, encouraging a pro-active regulatory style and the timely determination of issues. But it has also led to a proliferation of regulatory agencies, significant inconsistencies across industries, and a widespread feeling of lack of accountability.

A middle course would have to allow greater scope for regulatory initiative and discretion than is envisaged in New Zealand but a greater structuring of administrative discretion than in the United Kingdom. It would, in particular, more clearly limit the scope of regulation to areas in which there is a substantial danger of the abuse of market power; spell out in legislation the goals the
regulator is to pursue in these areas and the matters to be taken into account in making regulatory decisions; provide rules for the decisionmaking process; and allow a broader range of appeals mechanisms. It would also, I suggest, avoid establishing industry-specific regulatory bodies, in part for reasons of economy but in part to minimize the risk of inconsistency in the government process.

In all areas of public policy, the devil is in the details. This admonition is especially true in the regulation of natural monopolies, where policy aims at controlling powerful vested interests. There are no perfect ways of doing this, but our understanding of the legitimate goals, instruments, and processes involved has advanced greatly. Professor Jones and the organizers of this conference are to be commended for taking this opportunity to review the progress achieved to date.

Notes

1. I do not mean to suggest that governments ought to be indifferent to other markets. It is undoubtedly true that the self-correcting mechanisms that limit the abuse of market power will come into play more promptly where suppliers face fewer entry barriers, but does this mean that competition policy has as small a role to play as Professor Jones suggests? Rather, it seems reasonable to argue that cartels can be relatively durable—according to a recent OECD survey some cartels in Switzerland have operated successfully since World War I—and they are likely to be all the more so in economies with relatively few players and underdeveloped distribution systems. At a minimum, legislation should provide for private action against these cartels, preferably backed up by scope for class actions and punitive damages. Furthermore, a competition policy authority, in addition to providing for enforcement of prohibitions on collusion, should act as an advocate for competition in the economy, notably by identifying and drawing public attention to government measures that restrict the functioning of markets.

2. Last-offer arbitration is a dispute-resolution mechanism in which the arbitrator can only choose among the bids made by the parties (rather than splitting the difference). This usually encourages the parties to be more honest in their bids.

References


Jones agreed with Henry Ergas (discussant) that regulation is not a tea party—not in the OECD countries, and even less so in developing and former socialist countries, where information is scarcer and markets are more imperfect.

Much concern had been expressed at the conference about expropriation by the government of the earnings of the private sector. Jones observed that protecting the private sector from exploitation by the government was a valid concern, particularly in developing countries, but that people should be equally concerned about preventing the regulator from going to bed with the company and exploiting the consumer. In the case of the Philippines, which Brian Levy had discussed in an earlier session, Jones said the main reason for underinvestment was that the head of the telephone company was a friend of President Marcos's, and Marcos, who held shares under the table, was acting as a good monopolist would—maximizing profits by reducing quantity, which is done by underinvesting. An even better example was French Telecom from the 1950s to the 1970s. Jones understood that organization to be one of the least efficient telecommunications companies in the world. The system was difficult, costs were high, it took forever to get a phone and then it was hard to get a connection, and the quality of service was poor. Still, it was one of the more profitable private enterprises in France because by restricting investment the firm was able to charge very high prices.

Ashoka Mody (discussant in another session) elaborated on the problem of multiple providers in countries where competition is weak. Ghana, for example, has 50,000 telephones, probably only slightly more than the World Bank. This is not uncommon; many countries in Africa with populations of 15 million to 20 million have only 50,000 to 60,000 phones. They also have outdated technologies and extremely low call-completion rates, and during the day it is virtually
impossible to reach someone in the business district by phone. At the same time, new technologies—including mobile phones and satellite communication—are emerging that are going to radically change markets and public policy concerns in the next ten years.

The issue of competition must figure prominently in discussions about the leaps in technology that will be occurring, said Mody. Conservative estimates are that Ghana may need a million phones. India, which has 7 million phones, is expected to go to 100 million phones, at which point it will have lower market penetration than Malaysia has today. Consider, said Mody, the amount of investment needed in that sector and the range of opportunities in terms of technology and institutions. It is essential to think through the implications of multiple providers in a combination of geographic areas. Jones had found humor in the situation in Mongolia, where twenty-three companies provide service to fewer than 2 million people, but Finland has hundreds of companies, said Mody. In ten years small regions will probably have several companies providing telecommunications and, for that matter, electricity. Issues of regulation and competition must be addressed.

Anthony Churchill, speaker in another session, urged participants to think about an aspect of telecommunications that has received little attention: international regulation. Telecommunications is an industry, he said, in which monopolies are exchanging rents without any attention to the interests of consumers or of other broad groups.

Following up on Jones’s comments about capture, Levy said that the issues of capture and of administrative expropriation are essentially asymmetrical; basically, we are looking at a mirror image of the same problem rather than at two different problems. Both can take place down the road, within a regulatory system that creates flexibility and thus provides room for capture and expropriation. One way to resolve the problem of creeping administrative expropriation is through precise, specific rules—about substance or about process—that restrain administrative discretion. The same solution would apply to rules that restrain capture. Capture can take place at the time of the deal; the terms could be so favorable to the new buyer as to be essentially a capture. A transparent mechanism of privatization with open bidding would deal with the capture problem. Some might argue that is what happened in Argentina, although Levy and Spiller do not make that point explicitly, or even implicitly. How does one prevent capture that takes place at the time of the deal? The answer is to create a transparent mechanism of privatization with open bidding, the same way one avoids administrative expropriation.

Jean-Jacques Laffont, discussant in an earlier session, said Jones had compared distortion in pricing with distortion caused by defects in cost minimization. But Jones appeared not to have included in costs the enhancement given to private firms, which involves a public loss because of imperfections in the tax system. Should one include the costs related to the enhancement given to monopolies, and would this not be a serious problem in developing countries,
where tax systems are poor? Theoretically, if there are no such costs, the best course is to give the firm a transfer roughly equal to the social benefit produced by its activity.

Laffont was observing, said Jones, that if monopolists are simply allowed to practice perfect price discrimination, or if monopolies are given an appropriate equivalent transfer, they can be induced to operate at the efficient allocating point, providing that, as Laffont says, there is no cost to the transfer. That gets down to the question of the magnitude of the opportunity cost, or shadow multiplier, on government funds, as opposed to funds in private hands. The question is obviously correct, said Jones, if the multiplier on government funds is 1 or is the same as funds in private hands. One could argue that the multiplier on government funds is higher in less developed than in more developed countries, but that is a separate issue.

The real arguments for not going that route are not efficiency arguments. One rationale is that political opinion, whether in industrial, developing, or former socialist countries, would rebel at the idea that excessive profits were accruing to particular capitalists. To some extent, keeping prices and profits reasonable will preclude that kind of transfer. A second argument is that if such transfers were made, a large part would be chewed up in rent seeking and cost inefficiencies, given the control structure of firms in developing countries.