

1

Typical Justifications for Regulation

The major objectives of most economic regulatory efforts fall within one of the categories discussed in this chapter.¹ The *justification* for intervention arises out of an alleged inability of the marketplace to deal with particular structural problems. Of course, other rationales are mentioned in political debate, and the details of any program often reflect political force, not reasoned argument. Yet thoughtful justification is still needed when programs are evaluated, whether in a political forum or elsewhere. Usually this justification is one (or more) of the following.²

The Control of Monopoly Power

The most traditional and persistent rationale for governmental regulation of a firm's prices and profits is the existence of a "natural monopoly." Some industries, it is claimed, cannot efficiently support more than one firm. Electricity producers or local telephone companies find it progressively cheaper (up to a point) to supply extra units of electricity or telephone service. These "economies of scale" are sufficiently great so that unit costs of service would rise significantly if more than one firm supplied service in a particular area. Rather than have three connecting phone companies laying separate cables where one would do, it may be more efficient to grant one firm a monopoly subject to governmental regulation of its prices and profits. To understand why this may be so, we must examine the underlying arguments.

The Traditional Economic Rationale for Regulation

In a perfectly competitive market, firms expand output to the point where price equals incremental cost—the cost of producing an additional unit of their product.³ A monopolist, if unregulated, curtails production in order to raise prices. Higher prices mean less demand, but the monopolist willingly forgoes sales—to the extent that he can more than compensate for the lost revenue

from fewer sales) by gaining revenue through increased price on the units that are still sold.⁴ The result is waste: Consumers compare the high monopoly price of the monopolized product with the relatively cheaper prices of competitively produced products and buy more of the latter even though (1) they may prefer more of the former and (2) it costs society less in terms of real resources to produce more of the former and less of the latter.⁵ Thus, where economies of scale render competition wasteful, the classical economist or regulator will try to set price near incremental cost in order to induce the natural monopolist to expand its output to a socially preferred level—where buyers do not inefficiently substitute consumption of socially more costly goods for consumption of the monopolized good.⁶

Monopolists may also lack sufficient incentive to hold production costs at low levels. They do have some incentive to lower costs, since lower costs will increase their profits. But although there is a “carrot,” there is no “stick”: they do not feel the pressure of competitors who would threaten to lower their own costs, subsequently lower their prices, and thereby capture sales. For this reason, the monopolist may be lazy about production costs. The extent to which regulation can counteract this tendency is doubtful.⁷

Objections to the Traditional Economic Rationale

There are serious objections to the basic rationale for regulation just described. While none of them suggests outright rejection of the economic model, they do seriously temper the enthusiasm with which we embrace its results.

The “second best” problem. The theory of the “second best”⁸ casts doubt on the value of forcing prices down toward incremental costs. Under perfect competition, the price of each and every good reflects its incremental cost; buyers cannot be misled into choosing goods that they desire less but that cost society more to produce; resources are allocated so as to maximize social welfare. However, in the real world, differences between price and true economic cost not only exist, but vary greatly across goods and industries. The theory of the second best posits that one cannot readily determine what sort of price changes are needed to correct the resulting inefficiency. The prices of goods *relative* to one another, not their absolute level, direct purchasing decisions. Thus, monopoly may raise prices, but an equal degree of monopoly throughout the economy should lead to relative prices roughly similar to those of perfect competition. And it should give consumers much the same information about relative production costs as would perfect competition. In an economy in which there are varying competitive conditions and in which there is no accurate, practical method for measuring cost/price differences, one cannot be certain whether lowering the price of a single product will provide consumers with *more* or with *less* accurate information about the relative costs of products. Thus, one cannot be certain whether lower prices and increased production generated by regulation are more or less wasteful of society’s resources.

The “second best” argument weakens but does not destroy the classical rationale for regulation. One might confess to uncertainty but still believe that lower regulated prices are more likely to move the economy toward rather than away from allocative efficiency. Large sectors of the American economy are competitive—with prices that are presumably near incremental costs—and unregulated monopoly can raise prices far above costs. In addition, monopoly can significantly distort buying decisions within particular sectors in a way that is entirely independent of its effect on the economy as a whole. Assume, for example, that within the energy sector coal and natural gas are produced competitively but oil is monopolized. Relative prices induce consumers to use some coal that (in relative terms) is more socially costly to produce than equivalent oil. Increased competition in the oil industry improves allocational efficiency within the energy sector regardless of the effect it has on buying decisions elsewhere.⁹ In the absence of a strong empirical showing that such proregulation premises are untenable, one may find “second best” considerations ignored in regulatory debates for an important practical reason: agencies and courts need principles and standards to determine the substantive validity of regulatory decisions governing price and other (particularly anticompetitive) practices. The competitive model—based upon the desirability of incremental cost pricing—provides such standards. Given the difficulty of elaborating acceptable substitute standards, it is unlikely that it will be abandoned.

Price discrimination. Without regulation, natural monopolists will not necessarily restrict output, if they can discriminate in price.¹⁰ “Price discrimination” means charging, say, different customers two or more different prices for the identical product. For most producers price discrimination is impractical, because the customer who is charged a low price for a unit of the product can turn around and resell it to the customer whom the producer wishes to charge a high price. The very nature of telephone service, electricity, and natural gas, however, makes it difficult for a user to resell the service he receives. Thus, if the telephone company, electric producer, or natural gas distributing company can discern the relative value of the service to each customer (that is, the maximum price each would pay for the service), it can dramatically increase revenues by raising the price to those who will continue to use its service, while keeping prices low to prevent abandonment by others. There would then be none of the service curtailment that the classical economist/regulator fears the monopolist would bring about without regulation.¹¹

The force of this argument depends upon the ease with which the monopoly firm can in fact discriminate. To discriminate in price, customer by customer, is administratively impractical for these firms, as is price discrimination over individual units sold to a single customer. A practical discriminatory pricing system would divide customers into classes, and services into categories. Railroads, for example, charge higher rates to transport more valuable commodities; airlines, seeking to charge business travelers more, give discounts to those who can stay away for a week or longer; electric utilities have “business” and

“residential” rates. The unregulated profit-maximizing monopolist presumably would raise the prices of some services across the board within some classes and categories. Thus, to some extent, without regulation service will be curtailed. The seriousness of the classical “output curtailment” problem depends upon the practicality of determining appropriate price classifications, which in turn depends upon the firm’s knowledge of how demand responds to price changes. The less accurately the monopolist firm can estimate its customers’ response to price changes, the more likely that it would, if unregulated, impose price increases that lead to curtailed service.¹²

Not much market power. A related argument assumes that demand for natural monopoly services is fairly elastic—that is to say, a fairly small increase in price leads to a significant cut in demand for the product. If so, the monopolist will not make substantial price increases for fear that buyers will switch to other products. Alternatively, it may be fairly easy for new firms to enter the market to produce this or a related product, and higher prices may induce new (but less efficient) competitors to enter and attract customers. Thus, the natural monopolist—even without regulation—may be unlikely to raise prices very high, and, in the absence of regulation, production will not be curtailed significantly.¹³

But this argument and the “price discrimination” arguments raise several empirical questions: To what extent is price discrimination possible? How does one go about determining elasticity of demand for different sets of customers over different ranges of price? What is the height of entry barriers? Given the difficulty of answering these questions, it is not surprising that intuitive judgments about them differ when applied to particular industries—just as do estimates of the importance of “inadequate output” as a rationale for regulating monopoly.

The need to pay for investment. The classical “inadequate output” argument may be thought of as an argument for nationalization and not regulation. Consider Hotelling’s¹⁴ example of a bridge that costs \$25 million to build and will last virtually forever. Assume that the resource cost per crossing—the wear and tear on the bridge—is 50 cents. To charge more than 50 cents will reduce bridge use and prompt inefficient expenditures (such as driving an extra ten miles to avoid the toll). Yet if the bridge owner charges only 50 cents, how is he to pay back the investors who put up the \$25 million to build the bridge? The answer, the argument goes, is nationalization. If the government invests \$25 million, needed bridges are built, and if it charges a toll of only 50 cents once the bridge is built, no one is unnecessarily discouraged from using it.

Nationalization, however, has its own problems. Consider two common ones related to allocation and efficiency: (1) *How is the government to know when and where to build bridges?* Unless bridge users are prepared to pay not only 50 cents for wear and tear, but also enough additional money to pay the investment cost (whether through tolls or taxes), it is wasteful to build the

bridge.¹⁵ A private investor will build bridges only where users can, and will, pay enough to cover investment costs. The government can try to reproduce (or improve upon) the private investor’s decision by developing a cost/benefit calculus for each project and investing only in those that demonstrate adequate social returns. However, there is serious doubt as to whether such studies produce results as accurate as those flowing from the discipline imposed upon investors by the knowledge that users must in fact pay sufficient tolls if the investment is to be recovered.¹⁶ (2) *Are nationalized industries less efficiently operated than those run by private firms?* Do they run the risk of undue political interference? Ambrose Bierce, for example, defined a lighthouse as “a tall building on the seashore in which the government maintains a lamp and the friend of a politician.”¹⁷

If nationalization is rejected, then capital must be raised from private sources, and regulators must set prices that allow fixed investment—investment in rights of way or railway beds, representing unrepeatable expenditure—to be paid for and recovered. This need, in the case of a natural monopoly, may well lead to prices that exceed incremental costs. Thus, even when regulation works perfectly, it cannot, by setting prices equal to incremental costs, totally cure the “misallocation” that natural monopoly theoretically might cause, but rather it will aim at setting higher prices that allow recovery of investment costs when incremental cost pricing would not do so.

In sum, the traditional “inadequate output” justification for natural monopoly regulation rests upon the assumption that a natural monopoly, operating without fear of present or future legislation, would set prices not simply higher than incremental cost, but significantly higher. It assumes that the regulated prices (including the recovery of investment) generate product prices (in a sector or in the economy generally) that reflect relative resource costs better than would unregulated prices. Such a proposition is logical, but rests upon a host of empirical assumptions that are unproved or, as a practical matter, unprovable.

Additional Bases for Regulation

In addition to the economically based “increased output” rationale, three other justifications for regulating the natural monopoly are often advanced.

Income transfer. Even if the monopolist maintains output at roughly the competitive level—by practicing price discrimination, for example—he will raise the price of some of his services far above a regulated or competitive level. The effect is to transfer income from the users of the service to investors—an income transfer that is generally believed to be regressive, and hence undesirable.¹⁸ From the investor’s perspective, this transfer of wealth takes place only once. If, for example, an electricity company were suddenly deregulated and found it could earn profits equivalent to, say, 50 percent on its

equity investment instead of its previous earnings of 10 percent on the stock's market value, the price of the company's shares would rise, so that earnings on the new higher share value would be roughly equivalent to returns earned by other competing investments.¹⁹ Initial investors receive a windfall; future investors, who pay more for the shares, do not. Consumers indefinitely pay the extra charges each year that constitute the increased profits that the firm now enjoys. The more essential the service, the greater the amount that this income transfer is likely to be. Unlike monopoly rents, or patents, or rents accruing to firms that improve their products, there is no natural limitation to the time period over which increased profits are earned, nor is there any obvious social advantage in allowing the firm to earn them. Although the amount of probable income transfer in any specific case is almost always unknown, proponents of regulation may assume it to constitute a strong reason for regulating the prices and profits of the natural monopolist.

Fairness. The competitive market does not provide the firms within it much opportunity for the arbitrary or unjustifiably discriminatory exercise of personal power. If grocery store A hires rude salesmen or provides inadequate service, the customer can switch to store B. If an unregulated telephone company were to treat a customer unfairly, he or she would have no ready recourse. Given the high costs of litigation, courts are not readily accessible. Appeals to higher levels within the firm's bureaucracy may not be very effective when the firm is aware that the customer must continue buying the firm's service regardless of the outcome. The regulatory system, by providing recourse for grievances against the monopolist, offers a remedy that to some extent makes up for the lack of competition's guarantees against unjustified discrimination.²⁰

Power. Regulation is also advocated by those who fear concentration of substantial social or political power in the hands of a single firm that controls an essential product. Whether regulation is in fact effective in achieving this end is, of course, debatable. Yet the need to file reports, the fear that improper conduct might prompt hostile governmental action, the feeling of public responsibility that regulation may engender in the minds of the firm's executives—all suggest that regulation may have some effect (for better or worse) on the social or political activities of a large firm. Though the meaning and desirability of "corporate social responsibility" is controversial, it can be argued that regulated firms such as the telephone company will take actions popularly regarded as socially responsible, though there is no empirical data about whether regulation makes such actions more likely.

This survey of the classic case for economic regulation—regulating the prices and profits of the natural monopoly—although brief, suggests that the case for regulation rests partly on economic grounds, partly on political and social grounds, and upon a host of unproved (and possibly unprovable) assumptions. However, enough of these assumptions are plausible to make regulation an apparently reasonable governmental response to the natural monopoly.

Rent Control or "Excess Profits"

What Is a Rent?

Since "economic rent" is often confused with "monopoly profit," the problem of rent control is often confused with that of controlling monopoly power. Yet rents and monopoly profits are very different.

A firm will earn an economic rent if it controls a source of supply that is cheaper than the current market price.²¹ It is a rent and not a monopoly profit if the cheap source could not supply the entire market. If, for example, coal sells in a competitive market for \$20 a ton, and if Smith finds a small but unusually rich seam that can be mined for \$1 a ton, Smith will earn a rent of \$19 a ton. Smith's profit is best thought of as a rent because his output is limited. He cannot expand production to the point where he could supply the entire coal market—his seam is too small. Thus, if all non-Smith coal costs \$20 to produce, and buyers are willing to pay \$20, those buyers will bid the price of Smith's coal up to \$20.

Unlike a monopoly profit, the existence of a rent does not mean that there is "inefficiency" or "allocative waste." Assume that the coal industry is highly competitive and sells 15 billion tons per year—including 50 million from Smith's mine—at a price of \$20 per ton. By definition, Smith can produce no more; by definition, additional coal will cost \$20 or more per ton to produce. To sell coal at \$20 equates demand with supply and allows each buyer to see the additional resource cost of supplying his (extra) demand. To force the price of coal down below \$20 will yield no more coal, for it costs \$20 or more per ton to produce more than the existing 15 billion tons. In fact, a price below \$20 will bring about a shortage because customers will demand more coal and their demand cannot be satisfied at less than \$20. To lower price (below the resource cost of satisfying their demand) causes distortions in the prices of coal and related commodities.

It should be noted that rents are common throughout the economy, in competitive and noncompetitive industries alike. Any firm that finds a more efficient production process, that finds an unusually cheap supply source, that luckily buys a machine at a time when they are cheap, that has unusually effective managers—but that cannot expand to the point of satisfying a significant share of industry demand at prices that reflect its lower costs—earns a rent. To discourage the earning of rents is highly undesirable, for it would impede the search for efficiency. In many instances it seems perfectly fair that rents should accrue to producers who, through talents or skill, produced them.

The Rationale for Regulation

Many of those who demand regulation of rents may not be aware of the precise nature of the problem. They may see large producer profits and believe that they are monopoly profits or they may simply see prices rising, and, with-

out inquiry into the cause, exert political pressure to bring about lower prices through regulation.²² Mistake, confusion, or political power may cause regulation but cannot justify it. Yet a plausible justification for regulation to control some rents exists. The justification rests upon the desirability of transferring income—from producers to consumers—in a very few instances in which producer rents are large and occur suddenly.

Thus, one finds plausible arguments for rent control made where producers of certain products particularly important to consumers suddenly find that they, through no particular initiative of their own (through luck or change in general economic conditions), have earned very large rents.²³ During World Wars I and II owners of urban housing and property found that increased immigration into cities combined with the curtailment of new building (as construction workers were called to defense work) forced up the price of old housing.²⁴ Similarly, it has been argued that producers of natural gas that had been found at low cost could earn huge profits—by selling this old cheap gas in a free market, where increased demand and vastly increased exploration costs made natural gas far more expensive.²⁵ And, of course, between 1973 and 1980 the Arab cartel multiplied the value of existing oil stocks by a factor of ten.

These cases argue for regulation for reasons related not to more efficient use of the world's resources, but to a fairer income distribution. First, it is felt that the extra profit accruing to these producers is somehow undeserved. It does not reflect wise investment decisions—money was not attracted into oil exploration by anticipation of a future Arab boycott. Rather, it reflects plain luck. Second, the income transfer from consumers to producers or to their shareholders that these profits represent is thought to be *regressive*. Of course, the transfer will in part be taxed away—through tax rates on unearned income, for example, that approach 70 percent (compared with a 50 percent rate on earned income). But tax rates may fall, and producers may find ways within the tax laws to shelter the increased income from these high rates of tax. Third, the amounts involved are *large*—so large that the government should intervene to ensure that “windfall” rents are captured for the benefit of the consumer rather than the producer.

The claim for regulation is strengthened if the price of a product increases not just drastically but also *suddenly*, as did, for example, the price of oil during the Arab oil embargo. The uncontrolled market price of \$10–12 per barrel was five to six times the preembargo price; and the increase took place within a period of weeks. Such jumps obviously generated high rents for those controlling oil stocks. The rents were arguably “undeserved”; the transfer was *regressive*, and the amounts involved were large. Moreover, the *suddenness* of this rise might have required consumers to cut back other expenditures drastically in order to pay increased oil bills without significantly eroding their savings. The net result might have been to hasten economic recession.

In essence, rent control is aimed at transferring income and is undertaken for reasons of “economic fairness.” It may also seek to avoid certain adverse

social effects—dislocation and hardship—of a significant increase in the price of an essential household item. These hardships may result from the suddenness of the price increase rather than the existence of the rent.

Compensating for Spillovers (Externalities)

What Are Spillovers?

A considerable amount of regulation is justified on the ground that the unregulated price of a good does not reflect the true cost to society of producing that good. The differences between true social costs and unregulated price are “spillover” costs (or benefits)—usually referred to by economists as “externalities.”²⁶ If a train emits sparks that occasionally burn the crops of nearby farmers, the cost of destroyed crops is a spillover cost imposed upon the farmers by those who ship by train—so long as the shipper need not pay the farmer for the crop lost.²⁷ Similarly, if honeybees fertilize nearby apple orchards, the beekeepers provide a spillover benefit to the orchard owners—so long as the latter do not pay the former for their service.²⁸ Spillover benefits have sometimes been thought to justify government subsidy, as when free public education is argued to have societal benefits far exceeding the amount which students would willingly pay for its provision. Yet when one considers regulatory systems, spillover costs—not benefits—are ordinarily encountered.

The Classical Rationale for Regulation

Like the regulation of natural monopoly, the regulation of spillover costs is justified by the desirability of avoiding economic waste. Suppose a factory can produce sugar either through production method A or production method B. Method A costs 9 cents per unit of production but sends black smoke billowing throughout the neighborhood to the annoyance of neighbors for miles around. Method B costs 10 cents per unit of production and produces no smoke at all. The profit maximizing factory owner adopts A although, if those injured by the smoke would willingly pay more than 1 cent (per pound of sugar) to be rid of it, method A is socially more expensive. Then B, not A, should be chosen, because its total social costs—including costs of harm inflicted—are lower. As long as the affected public prefers reduced pollution to its noisome effects, it should bribe the producer to choose method B. Where the public prefers reduced pollution yet finds no practical way to bribe the producer, too many of society's resources are attracted (by lower prices not reflecting the cost of pollution) into polluting processes and products, and too few are attracted into pollution-free products and processes. Government intervention arguably is required to help eliminate this waste.

Objections to the Classical Rationale

It can be argued that spillover costs do not call for government intervention but, rather, for a rearrangement of private property rights.²⁹

First, as Ronald Coase³⁰ has pointed out, if bargaining were costless, spillovers would not exist. Those suffering the pollution in the example above would simply band together and offer to pay the factory to use process B rather than A. The factory owner would switch only where the coalition was willing to pay more than 1 cent per pound. Precisely the same result would occur if the factory were required to compensate residents for pollution damage (both physical and psychological) it created, or if the government intervened and stopped pollution precisely and only in each instance where the public would rather pay abatement costs than suffer the effects of pollution. Thus, one might argue, why not let people bargain privately to abate pollution rather than introduce government regulation?

The answer to this question is that bargaining is not costless. Thus, the residents may suffer the pollution despite a willingness to pay more than 1 cent per pound to avoid it, simply because it is too difficult for them to band together. As the number of affected people increases, communication becomes more expensive, bargaining becomes more complicated, and a clear consensus is harder to obtain.³¹ Furthermore, there is the added problem that some participants may systematically underestimate the true value of abatement to them in the hope of minimizing their contribution to the cost of abatement.³² Thus, transaction costs may permit the continuance of spillover costs even though society would be better off without them.

Second, one may object to the classical rationale for regulation on the grounds that inefficient spillovers could still be eliminated by allocating relevant property rights to those who are most likely to achieve the efficient result through bargaining (were bargaining practical). For example, making the factory liable for the damage it causes will prompt it to switch to process B where doing so is cheaper than compensating sufferers. Freeing the factory of liability places the burden of organizing on the sufferers, so that they can bribe the producer to switch to B—an impractical alternative.

Third, one might argue that rather than introducing governmental regulation, it might be better to create liability rules to allocate the burden of paying for the harm of pollution in a way that leads to "efficient" solutions. However, court-administrated liability rules, although they may indeed have theoretical and practical merit for dealing with some spillovers, cannot adequately deal with others. For one thing, government officials must determine the precise shape and location of liability rules. This in itself may prove a herculean task. For example, should liability be placed upon the party best able to calculate the relevant costs?³³ Upon the party best able to organize? Upon the party most likely to reach the socially optimal result? It may be as difficult for a "liability rule maker" to determine the correct response in a particular instance as for a

central administrator to choose between process A and process B. Moreover, the court enforcement of liability rules may be nonuniform, expensive, and have as many harmful side effects as central administrative direction. Thus, despite the theoretical existence of a system of private rights to deal with spillovers, direct governmental intervention is often believed desirable.

Another, less serious objection to governmental regulation rests upon the belief that spillovers do not exist. Undesired air pollution, for example, is caused not only by a process that sends smoke up the stack but also by the presence of nearby residents who breathe the smoke. The discomfort is as much a result of living near the factory as of using process A: while the cost is external to the good produced by A, it is internal to those who live near the factory and affects the behavior of those thinking of moving. And moving away eliminates the problem just as would the factory's switch to process B.³⁴

This objection, however, misses the point. As a technical matter, externalities (or spillovers) are defined in relation to particular products. That is to say, air pollution is spillover in relation to the production of sugar; it is not claimed to be a spillover in relation to living near a sugar factory. Moreover, the justification for governmental intervention to deal with this spillover rests upon the judgment that society is better off where firms are encouraged to control polluting processes than where persons are encouraged not to live near polluting factories. Insofar as this justification relates to the spillover, it reflects the belief that those affected by air or water pollution would rather pay the cost of reducing the pollution than suffer from the pollution itself. Furthermore, the difficulties of organizing coalitions to bribe producers are just too great under the present allocation of private rights and liabilities. Thus, governmental intervention is likely to approximate better the amount of pollution that affected consumers are willing to pay for.

Of course, intervention itself is not costless. Moreover, intervention—or rearrangement of rights and liabilities—changes the distribution of wealth and income. Those who buy sugar (and the owners of sugar factories) are made poorer and those who suffer pollution are made richer by intervention or liability adjustments leading to reduced pollution. Insofar as intervention is designed to produce the result that would be paid for were bargaining feasible, those suffering pollution will gain more than the others lost. Whether this is a sufficient basis for intervening or altering liabilities may be debated. Traditionally it has been argued that changes should be made when the beneficiaries can compensate the losers out of their gains and have some gain left over; others have claimed that since the compensating is only hypothetical and not actually carried out, there is no clear justification for making the change. Regardless of the merits of these arguments, it is sufficient to note here that this rationale is one of "economic efficiency." To satisfy it is to move closer to a world in which all resources are used in a manner that maximizes the welfare of the world's individuals as measured by their preferences revealed in the marketplace. And it is the same rationale of allocative efficiency that underlies many

governmental economic decisions and actions, including regulation of natural monopolies.

In sum, a spillover rationale must be phrased in terms of a particular product; it must assume that obstacles to bargaining lead to significantly greater use of a product (or production process) than would be the case if costless bargaining were possible; and it must assume that the result of intervention (taking into account the costs of intervention) will better approximate the bargained-for solution. If these assumptions are correct, then intervention will reduce allocative inefficiency.

A Caveat

One can too readily classify as a spillover cost or benefit almost any policy reason for taking actions that lead to results other than those dictated by existing market arrangements. Thus, law enforcement can be seen as securing the "external benefits" of security, or the laws prohibiting blood feuds can be justified as elimination of the "external costs" of chaos and disruption. Where it is used to describe those commodities, the value of which is incapable of even rough monetary estimation—commodities such as justice, security, and so on—the "spillover" notion is virtually useless. The more readily monetary values of the commodities or objectives can be estimated, the more directly useful "spillover" characterization is likely to prove. In other cases, one is better off speaking directly of noneconomic reasons for and against taking a particular action rather than explicitly invoking the notion of "spillover."

Moreover, since there is always some possible beneficial effect in reversing a market-made decision, one can always find some (broadly defined) spillover cost rationale for regulating anything. Thus, the rationale, if it is to be intellectually useful, should be confined to instances where the spillover is large, fairly concrete, and roughly monetizable.

Inadequate Information

For a competitive market to function well, buyers must have sufficient information to evaluate competing products.³⁵ They must identify the range of buying alternatives and understand the characteristics of the buying choices they confront. At the same time, information is a commodity that society must spend resources to produce. The buyer, looking for alternative suppliers, spends time, effort, and money in his search.³⁶ The seller spends money on research, labeling, and advertising to make his identity and his product's qualities known. In well-functioning markets, one would expect to find as much information available as consumers are willing to pay for in order to lower the cost or to improve the quality of their choices.

The Classical Rationale for Regulation

Markets for information may on occasion not function well for several reasons. First, the incentives to produce and to disseminate information may be skewed. Like the bridge in Hotelling's example, some information (particularly that requiring detailed research) is expensive to produce initially but very cheap to make available once produced. Since it can be repeated by word of mouth, televised, or printed and reprinted at low cost, it may easily benefit many recipients who never pay its original producer. Thus, those in the best position to produce the information may not do so, or they may hesitate to disseminate it, for fear that the benefits will go not to themselves but only to others.

The importance of this problem varies considerably depending upon the type of information and its use. A firm that manufactures breakfast foods, for example, would have every incentive to produce information showing that its cereal was more nutritious than that of its competitors and to disseminate that information widely. Moreover, the production, use, and dissemination of much information is protected by copyright and patent laws.³⁷ Further, the inadequate incentive to produce information typically leads to a demand not for regulation but for governmental support of production and dissemination.

Nonetheless, occasionally the problem may lead to a demand for regulation. Drug manufacturers, for example, are required to print the generic (general scientific) name of their product, as well as the brand name, on the label. Thus, the buyer sees that a host of competitors in fact offer to sell the same product. This labeling requirement can be seen as lowering the cost to buyers of searching for competing sellers,³⁸ by quickly making them aware of the competitors' existence. And it does so by requiring those with the information most readily at hand to make it available.

Second, one of the parties to a transaction may seek deliberately to mislead the other, by conveying false information or by omitting key facts. A seller of securities may lie about the assets of the company; a seller of a used car may turn back the mileage indicator. Of course, false statements or active misrepresentations may be grounds for rescinding a contract or suing for damages. Yet the cost of court action is often high enough to weaken it or give it minimal effect as a deterrent. Nor can one necessarily rely upon fear of declining reputation to act as a deterrent. The importance of reputation in securing sales depends upon the particular product, the particular seller, and a host of other circumstances. The rationale for governmental action to prevent false or misleading information rests upon the assumption that court remedies and competitive pressures are not adequate to provide the consumer with the true information he would willingly pay for. Thus, the Securities and Exchange Commission (SEC) regulates the issuances of securities, while the buyer of used cars is typically left to his basic judicial remedies.

Third, even after locating potentially competing sellers, the buyer may not be able to evaluate the characteristics of the products or services they offer.

The layman cannot readily evaluate the competence of a doctor or lawyer. Nor can he, unaided, evaluate the potential effectiveness or dangers of a drug. And he is unlikely at the time of purchase to know if a car is a lemon. Formal or informal understandings among those on the supply side—whether doctors, lawyers, or drug producers—may make difficult or impossible the creation of objectively applied labels that aid evaluation.³⁹ Governmental intervention may be desired to prescribe the type of information that must be provided, as well as to help buyers evaluate the information that is being supplied.⁴⁰

Fourth, the market may, on the supply side, be insufficiently competitive to provide all the information consumers would willingly pay for. Until the government required disclosure, accurate information was unavailable to most buyers concerning the durability of light bulbs, nicotine content of cigarettes, fuel economy for cars, or care requirements for textiles.⁴¹ In the 1930s automobile manufacturers advertised the comparative safety of their product. Subsequently this advertising disappeared, since auto makers felt that calling attention to safety problems hurt the industry more than it benefited individual firms. For similar reasons one does not find individual airlines advertising safety records. Since the airline industry is highly competitive in many respects, this fact suggests that tacit understandings not to supply certain varieties of information may be easier to reach (the industry need not be highly concentrated) than are tacit agreements not to compete in price or in service quality.

Criticisms of the Rationale

Criticisms of the rationale for regulating the provision of information usually focus on whether the rationale applies to the particular case at issue. Critics may claim, for example, that in a particular case the market is functioning competitively, consumers are sufficiently capable of evaluating a product's qualities, or there is little deliberate deception. They may argue that a particular agency's efforts to provide information are too expensive, that the information is unnecessary, that disclosure itself may mislead consumers, or that it may interfere with the competitive workings of the marketplace. For example, the efforts of the Federal Trade Commission (FTC) to require that imported products have a "country of origin" label have been criticized as at best a waste of time and at worst as an effort to protect American manufacturers from foreign competition by imposing costly labeling requirements on foreigners.

In sum, there is little quarrel with governmental efforts to help consumers obtain necessary information when the information is in fact needed and the intervention lowers the cost of providing it. Critics of intervention tend in particular cases to quarrel with the claim that regulation will lower the costs of its provision.

Excessive Competition: The Empty Box

A commonly advanced rationale for regulation of airlines, trucks, and ships is that competition in those industries would otherwise prove "excessive." This rationale has been much criticized as incoherent or at least inapplicable to the transportation regulation it is meant to justify. In fact, the difficulty with the term is that it has been used to describe several different types of rationale—some of which are no longer acceptable justifications for regulation. The notion common to all those rationales is that prices, set at unprofitably low levels, will force firms out of business and result in products that are too costly.

Historical Use

The history of airline and trucking regulation offers some insight into two possible uses of the term "excessive competition."

Airlines. The federal government began to regulate prices and profits of the airline industry in the 1930s, when it was still subsidizing the industry to encourage its development. The subsidy consisted of contracts to carry airmail at prices that exceeded the cost of carriage. A major scandal arising out of subsidy awards by the United States Postal Service led Congress, first, to transfer subsidy authority to the Interstate Commerce Commission (ICC), and, second, to consider comprehensive change of the subsidy system. The ICC administered the subsidy by asking the airlines to bid for the right to carry U.S. airmail. It awarded the contract to whichever line was willing to fly for the lowest subsidy. But the airlines believed that once the government awarded a contract to a particular carrier, it would increase the subsidy should the initial payment prove insufficient to cover the airline's costs. As Colonel Gorell, an industry representative, testified before Congress in 1937, "The law put a premium upon an unreasonably low bid since there is always the possibility that later on, a rate first put unjustifiably low will be raised by the action of the Interstate Commerce Commission . . . [which] is ultimately under the duty of fixing a reasonable rate."⁴²

As a factual matter, it is not clear that competition in the 1930s was "excessive" in any sense. Colonel Gorell also testified before Congress in 1937 that excessive competition had not actually occurred, but that it had been "much closer than I would like to talk about."⁴³ Moreover, the classical price and profit regulatory provisions of the act may reflect the fact that its framers borrowed many of its provisions from the act governing the ICC.⁴⁴ But, at least in theory, a subsidy offer provides one perfectly sensible rationale for "excessive competition": the airlines' belief that the ICC would increase the subsidy award to a firm under contract provided an incentive to bid for the initial contract at a price well below cost. By doing so, an airline could expand the size of its route system, with the government making up the difference. Indeed, each airline would charge low prices to all customers, for the objective

of each would be to expand system size and not to earn profits or to minimize losses. Thus, a government seeking to minimize the amount of subsidy required would have to prevent prices that were "unreasonably low," and regulation would be justified in order to minimize government outlay.⁴⁵

This rationale for regulation, while coherent, is not applicable today, when almost all airlines are unsubsidized.

Trucks. Demand for regulation of trucking prices, profits, and entry arose in the 1930s for three reasons. First, the railroads, regulated by the ICC, complained that truckers were undercutting their prices. They did not argue that trucking prices were below costs, but rather that the ICC required the railroads to charge more than incremental costs on routes where they could compete with trucks; thus, the trucks took away business that, on a "least cost" basis, should have belonged to the railroads. This argument does not offer a strong rationale for regulation of trucking, but rather suggests the need for changing the way railroads are regulated.

Second, the truckers argued that an unregulated market would lead some firms to cut prices and drive others out of business. Regulation would keep more firms in business and provide more employment. In principle, however, competition drives firms out of business because the survivors can do the same job better, more efficiently, or with fewer employees. To keep *unnecessary* firms in business is likely to sound more reasonable at a time of very serious depression (like the 1930s) than today.

Third, in the 1930s the competitive process itself was often blamed for the Depression. The framers of the National Industrial Recovery Act, for example, believed that agreements among firms not to cut prices would increase profits, encourage investment, and maintain employer purchasing power. Thus, ordinary competition in many industries was viewed as "excessive." The NIRA cure for recessions, however, has been discredited. In a world in which competition is the desired norm and regulation the exception, the NIRA theory does not provide a coherent rationale for selective regulation.⁴⁶

Current Use

Currently, "excessive competition" might be used to refer to any of three alleged justifications for regulation.

The "natural monopoly." The "excessive competition" argument would make sense as applied to an industry that is a natural monopoly. One might claim that, without regulation, too many firms would seek to enter, and the resulting fight for market share would lead to the demise of all but one—the one that obtained the lion's share of the market by achieving the lowest unit costs.⁴⁷ According to the argument, this competitive process is wasteful of resources; a regulator should make certain that only one firm enters the industry and others do not seek to displace it.

Those opposed to regulation argue the opposite. They claim that any poten-

tial waste is justified by (1) the ability of the competitive process to pick the "best" firm out of the contenders; and (2) the tendency of the competitive battle to demonstrate empirically whether or not the industry is in fact a natural monopoly.

Regardless of the outcome of this argument, it does not apply to airlines, trucks, or ships, because those industries are generally conceded to be structurally competitive—not natural monopolies. Yet it is in the case of airline, trucking, and ocean shipping regulation that one finds the excessive competition argument used.

The cyclical nature of demand. "Excessive competition" might refer to a claim that, unless a particular industry is regulated, the cyclical nature of demand for its product will produce waste. When demand falls during, say, an economic downswing, competition among firms results in prices that cover only short-run incremental costs and prompts capacity curtailment. When the next upswing occurs, a firm will have to reopen its plant or rebuild its capacity. It might be argued that, rather than have firms go through the expensive process of closing and reopening plants, the government should set minimum prices—or allow firms to agree on minimums. Prices would remain high enough to cover fixed costs and plants would continue to operate during the recession. The extra costs of the agreements to consumers, one might claim, would be less than the cost of closing and reopening during the business cycle.

This argument assumes, however, that firms cannot raise sufficient funds in the capital markets to support excess capacity during periods of soft demand. If capital markets are functioning well, the firm ought to be able to attract funds to keep its plant open during the downswing on the basis of the expected future profits in the upswing. If it cannot do so, perhaps the capacity ought to close permanently, for consumers will not pay enough for the product (in the upswing) to justify maintaining it (in the downswing). Only if markets fail accurately to reflect the firm's future earning ability is it desirable to keep the plants operational.⁴⁸

Moreover, to prevent excessive competition by maintaining higher-than-competitive downswing prices may easily encourage overinvestment. Firms would be prevented from the otherwise unprofitable consequence of overpredicting demand for the industry and product. Any resulting overinvestment is a wasteful byproduct of an effort to prevent "waste."

In any event, this excessive competition argument applies to industries with large fixed investments and comparatively small variable costs—industries such as steel or copper.⁴⁹ One can understand that a recession might drive copper prices so low that mines must close, that the expense of reopening them may be great, and that capital markets may be unwilling to finance them so that they could remain open. But the argument does not apply to trucking, where fixed costs are low, or to airlines, where fixed costs can be reduced by leasing or storing planes during the downswing in the cycle.

Predatory pricing. Another variation of the excessive competition argument

is that a particular industry is subject to "predatory pricing." Competition leads some firms to price below costs, driving their competitors out of business. The remaining firm then raises its prices to excessive levels, leaving the public worse off than before.

For a firm to have an incentive to price predatorily, however, two preconditions must be met: (1) the predator must be powerful enough to outlast its competitors once prices are cut below variable costs; and (2) reentry into the market must be so difficult that the predator can maintain prices well above costs long enough to recoup its prior losses. Unless a firm is reasonably certain that both these conditions will obtain, it is irrational for it to attempt predatory pricing.⁵⁰

For this reason, it seems unlikely that predatory pricing will ever justify regulation. In fact, regulation can make predatory pricing easier, since it often provides the barriers to entry necessary for a potential predatory pricer to succeed. Furthermore, the antitrust laws make predatory pricing unlawful.⁵¹ Those firms suffering its consequences can bring antitrust suits and appeal to enforcement agencies. Of course, the enforcement agencies may not be effective, but that is no argument for regulation since the regulatory agency is as likely to be ineffective.

Moreover, unfortunately, ordinary price competition is easily confused with predatory pricing. The former generally involves low-cost firms charging lower prices that take business from higher-cost firms;⁵² the latter involves short-term prices well below costs, set with the object of destroying competition and later recouping losses through prices well above cost. Those advocating regulation on these grounds in the transportation field may well have confused the two.

Other Justifications

The reader should be aware of several other possible justifications for regulatory systems. While important, they have been used less often in the United States than elsewhere to justify governmental regulation of individual firms.

Unequal bargaining power. The assumption that the "best" or most efficient allocation of resources is achieved by free-market forces rests in part upon an assumption that there is a "proper" allocation of bargaining power among the parties affected. Where the existing division of such bargaining power is "unequal," it may be thought that regulation is justified in order to achieve a better balance. It is sometimes argued, for instance, that the "unequal bargaining power" of small sellers requires special legislative protection. While in principle one might regulate the "monopoly buyer" in order to protect these sellers, the more usual congressional response is to grant an exemption from the antitrust laws, thus allowing the sellers to organize in order to deal more effectively with the buyer. This rationale underlies the exemption granted not only to labor, but also to agricultural and fishing cooperatives.⁵³

Rationalization. Occasionally governmental intervention is justified on the ground that, without it, firms in an industry would remain too small or would lack sufficient organization to produce their product efficiently.⁵⁴ One would ordinarily expect such firms to grow or to cooperate through agreement, and to lower unit costs.⁵⁵ But social or political factors may counteract this tendency.⁵⁶ In such circumstances, agencies have sought to engage in industry-wide "planning." In the 1960s, for example, the Federal Power Commission argued that increased coordination in the planning and operation of electric power generation and transmission facilities would significantly lower unit costs. The commission felt that environmental, political, regulatory, and managerial problems make it difficult for firms to plan jointly. The result was a relatively unsuccessful federal agency effort to encourage industrywide rationalization.⁵⁷

Moral hazard. The term "moral hazard" is used to describe a situation in which someone other than a buyer pays for the buyer's purchase.⁵⁸ The buyer feels no pocketbook constraint, and will purchase a good oblivious to the resource costs he imposes upon the economy.⁵⁹ When ethical or other institutional constraints or direct supervision by the payer fail to control purchases, government regulation may be demanded.

The most obvious current example is escalating medical costs.⁶⁰ As medical care is purchased to an ever greater extent by the government or by large private insurers (with virtually no constraint on the amount demanded by the individual users), medical costs have accounted for an ever greater proportion of the national product.⁶¹ The fact that purchases are paid for by others frees the individual from the need to consider that using more medical care means less production of other goods; thus, he may "unnecessarily" or "excessively" use medical resources. If one believed that too much of the gross national product is accounted for by medical treatment, and also believed that the problem of moral hazard prevents higher prices from acting as a check on individual demand for those resources (which in turn reduces incentive to hold down prices), one might advocate regulation to keep prices down, improve efficiency, or limit the supply of medical treatment.⁶²

Paternalism. Although in some cases full and adequate information is available to decision makers in the marketplace, some argue that they nevertheless make irrational decisions and that therefore governmental regulation is needed. This justification is pure paternalism: the government supposedly knows better than individuals what they want or what is good for them. Such distrust of the ability of the purchaser to choose may be based on the alleged inability of the lay person to evaluate the information, as in the case of purchasing professional services, or the belief that, although the information could be accurately evaluated by the lay person, irrational human tendencies prevent this. The latter may be the case where small probabilities are involved, such as small risks of injury, or where matters of life and death are implicated, such as when those suffering from cancer will purchase a drug even though all reasonably reliable

information indicates that it is worthless or even harmful. Whether the brand of paternalism based on mistrust of consumer rationality is consistent with the notions of freedom of choice that underlie the free market is questionable. However, it plays an important role in many governmental decisions.

Scarcity. Regulation is sometimes justified in terms of scarcity.⁶³ Regulation on the basis of this justification reflects a deliberate decision to abandon the market, because shortages or scarcity normally can be alleviated without regulation by allowing prices to rise. Nonetheless, one might decide to abandon price as an allocator in favor of using regulatory allocation to achieve a set of (often unspecified) "public interest" objectives, such as in the case of licensing television stations. Sometimes regulatory allocation is undertaken because of sudden supply failures: to rely on price might work too serious a hardship on many users who could not afford to pay the resulting dramatic price increases, as in the case of the Arab oil boycott. "Scarcity" or "shortage" calling for regulation may also be the result of the workings of an ongoing regulatory program, as when natural gas must be allocated because of rent control or when an agency awards licenses to enter an industry.

The Mixture of Rationales

Many existing regulatory programs rest upon not one but several different rationales. Thus, for example, one might favor regulation of workplace safety for several reasons. One might believe that employers and employees can bargain fairly and equally for improved workplace safety (greater safety expenditures), but argue that accidents impose costs on others who are not represented at the bargaining table; thus, bargaining alone will produce inadequate expenditures for safety devices. This is a *spillover* rationale. Or one might believe that the worker does not know enough about the risks or consequences of accidents, so that he will fail to insist upon adequate safety expenditures. This is to argue that there is an *informational defect* in the market. Or one might feel that the worker is too poor or too weak to bargain for the safety he needs—that he has *unequal bargaining power*. Finally, one might claim that workers (indeed, all people) are simply incapable of understanding their likely future feelings about accidents that hurt them. They inevitably underestimate the risk. If regulation is an effort to give them what they "really" want (contrary to their expressed views), a *paternalistic* rationale is at work.

The importance of distinguishing rationales lies in the extent to which different rationales may suggest different remedies. Thus, one who believes that the primary problem is informational will tend to favor not classical regulation, but governmental efforts to provide more information. Although one who accepts a paternalistic rationale may disagree with one who believes the problem is informational, the clear statement of their points of difference can form the basis of empirical work that will lead them toward agreement upon the basic

rationale and thus help choose the regulatory weapon best suited to the problem at hand.

Similarly, the debate over the need for regulation of medical costs might be clarified if its proponents specified the rationale, or mixture of rationales, for regulation and the relative importance of each. To point to the increased price of medical care does not, by itself, suggest a need for regulation. The increased price might reflect cost increases due, for example, to medical advances. If scientific progress means that few older people die of pneumonia, more will (eventually) die of cancer or strokes, which require more expensive care and treatment. Moreover, labor costs may be increasing, as well as the costs of technology. Demand may increase because people have more money to spend on medical care. If rising prices reflect no more than increased demand for medical care (greater relative desire for medical care compared with other goods, greater ability to pay for it) and increased costs of supply (more technology, higher labor costs), regulation of these prices would not rest upon a "market failure" rationale. Indeed, rising prices might be a consequence of highly desirable actions taken by governments on grounds of equity—namely, supporting medical care for those who cannot readily afford it.

The proponent of regulation might cite other factors, however. He might argue that there is excess demand because so many patients do not pay their own bills—a problem of *moral hazard*. He might point to the difficulty potential patients have in determining whether they need care or what sort they may need. He might add that doctors themselves may not fully comprehend the economic costs of the treatment choices they make and thus choose treatment that is too expensive. All these are *informational problems*. Finally, he might fear that sudden increases in demand for medical care will lead to higher profits for hospitals, which, being "nonprofit" institutions, invest the "excess" in new, more expensive technology and plants. This is a problem akin to *rent control*.

Again, a breakdown by rationale does not determine whether hospital prices ought to be regulated. To know that, one would have to obtain empirical confirmation that the rationale is empirically important. Nonetheless, analysis may help clarify and focus the debate and thereby help policy makers reach more sensible conclusions.

This chapter has surveyed the major economic rationales for regulatory programs. Individually or in combination they underlie most major regulatory programs, which are themselves of several different types. The next six chapters each describe one major type of regulation. Together with this chapter, they provide a survey of most major types of regulatory programs and the economic justifications that underlie them.