INCREMENTAL COSTS FOR INCREMENTAL DECISIONS

IN TELECOMMUNICATIONS

By

Sanford V. Berg*
And
Dennis L. Weisman*

August 29, 1991**

Diversification by incumbent telecommunications companies raises many issues for regulators. Given the concerns expressed by potential competitors and some consumer groups, it is important for regulatory commissions to adopt costing principles that cannot be manipulated by participants in the hearing process. Incremental cost is offered as the benchmark concept for identifying predation and cross-subsidization in that it dutifully embodies the principle of cost-causality. Moreover, while it is important to measure costs with precision, it is better to have a rough approximation of the right number than a precise calculation of the wrong number.

* Sanford V. Berg is Florida Public Utilities Professor, and Dennis L. Weisman is Research Fellow, Public Utility Research Center, University of Florida, Gainesville, Florida 32611, USA.

In a recent article appearing in this Journal, Meitzen shows how regulatory cost allocation procedures can distort pricing and investment decisions.[1] They can lead telecommunications companies to choose separate subsidiaries rather than subject themselves to "accounting safeguards." Although these rules purport to protect customers from abusive market power and to ensure fairness for suppliers of competitive services, the resulting lost economies of scope represent a potentially large cost of regulation. If we adopt Meitzen's recommendation and shift from cost-based to price-based regulation for monopoly services, there is still a need to agree on costing principles which support economic efficiency, and permit the identification of potentially abusive behavior by incumbents, particularly cross-subsidization and predation. If key stakeholders in the regulatory process doubt the legitimacy of costing principles utilized therein, we are doomed to more lengthy and complex regulatory proceedings than ever before.

Of course, the issues are more complicated during this transition to competition and selective deregulation in telecommunications. In moving toward a competitive industry, both potential entrants and incumbents want to be sure that regulatory rules do not put them at a competitive disadvantage. Similarly, customers have new options (and risks), so they, too, have a stake in the design of rules. The regulatory process becomes entangled with intervenors: the hearing room becomes the place where market advantage can be won or lost. In some sense, if a potential entrant can be beat in the hearing room by being denied access to essential facilities, there is no need to beat them in the marketplace.[2] Similarly, if the incumbent is handicapped with burdensome rules and regulations, it may be unable to compete for some businesses which economically, it would be most efficient to serve. And, in fact, it may be easier for interested parties to "win" in the hearing room than in the marketplace.

One of the more controversial and complex regulatory issues today is the ongoing debate over costing principles and methodologies, since these might dictate the outcome of regulatory

investigations. Indeed, the evolution of the role of costing in telecommunications is an interesting and tortured one. According to Thomas McCraw[3], Alfred Kahn, as head of the New York Public Service Commission, was one of the first regulators to require that telephone companies submit cost data to support their rate design requests. At the time, New York Telephone had virtually no personnel dedicated to developing costs for service offerings. Today, the costing organizations in each of the seven Bell Operating Companies are large entities with many employees; in addition, numerous professionals work on issues involving jurisdictional separations.

Today, it is expected that costs should serve an instrumental role in regulatory proceedings. But there are many pitfalls in development and meaningful interpretation of cost data. Indeed, as discussed below, some costing methodologies are grounded in sound economic theory and decision science, while others represent little more than arbitrary manipulation of data designed to rationalize some predetermined outcome. Yet the costing issue itself has proved somewhat intractable in the academic literature, so it is little wonder that confusion abounds in the hearing room with expert witness poised against expert witness and vested interest against vested interest. Indeed, as Professor John Wenders has pointed out, the costing docket at the FCC went on for some 30 years--absent resolution as to the appropriate methodology.[4] (Note that in a similar vein, it took the FCC some 15 years to rule whether AT&T's Tel-Pac tariffs were legal).[5]

This paper discusses a set of costing principles appropriate for regulatory commissions and telecommunications companies in the Information Age. This overview does not provide a comprehensive treatment of costing methodologies, but sketches some key issues and principles. Perhaps the most complete set of investigations to date appears in *Marginal Cost Techniques for Telephone Services: Symposium Proceedings*, a National Regulatory Research Institute publication edited by William Pollard, January 1991. This volume contains papers by some of the leading experts in telephone costing. Pollard identifies three fundamental insights[6]:

1. *There is no one correct measure of marginal costs applicable to all policy objectives of a commission, company or court. This fact is not much comfort for a commissioner or policy maker desiring a single cost estimate on which to base all policy decisions. However, long-run incremental costs measured over a company's planning horizon may minimize strategic behavior on the part of participants to a regulatory proceeding because the cost of capacity is included.*

2. *Engineering process models are superior to econometric and optimization models when marginal-cost information is desired on a functional basis (that is, switching) or for a service such as local.*

3. *The terminology used to describe marginal-cost estimates in the economic literature and used by consultants is ambiguous and confusing, particularly to the noneconomist. Consequently, even the adoption by a commission of an incremental cost standard, whether long run or short, run may be insufficient to avoid confusion and unnecessary debate.*

We do not attempt to summarize the NRRI volume, but rather, focus on Pollard's first insight. In the process of presenting the case *for* incremental costs, we try to minimize terminological confusion noted in the third insight.

Regulatory Interest in Costing Principles

Why should regulators be interested in costing principles? The answer is complex and multi-dimensional. If we were to list reasons, we might present the following:

1. **Achievement of Economic Efficiency**

   The benchmark criterion for economic efficiency is a competitive marketplace in which prices are driven to the vicinity of marginal or incremental cost and excess profits are zero.

   To this point in our discussion, we have used the term "cost" rather loosely, without defining

it precisely. When the economist uses the term "cost," it generally means some measure of *marginal* or *incremental* cost. That is, for efficiency, prices should be set in such a way as to reflect the true resource costs borne by society in producing the marginal or incremental unit of the good of service. A price set higher than marginal cost will preclude some market transactions that could have made both the firm and the consumer "better-off." A price set lower than marginal cost will induce overconsumption--some agents will purchase the service when their valuation of the service is less than the resource cost borne by society in producing it. Both of these situations are sub-optimal. Hence, from an efficiency perspective, so-called first-best resource allocation is obtained when prices are set equal to incremental cost, assuming there are no market failures (like external effects) or market imperfections (like inadequate information for consumers).[7]

2. Concern For Fairness/Income Distribution

The regulator, as a "social planner," maintains a de facto taxing authority by setting rates for services consumed by the vast majority of the population. Of critical concern in this capacity are the fairness and income distributional impacts associated with the choice of price structures for customer classes and product lines.

Within the scope of the regulators' authority is the power to establish public policy in a manner than best serves the *public interest.*[8] For example, regulators may decide on the basis of fairness and the collective social good that certain services, e.g. basic local service, should be priced so as to ensure that every individual desiring such service shall not be denied access because of inability to pay. This dimension of fairness involves an *economic right* to a commodity (such as universal telephone service). The particular service needs to be well defined and the sources of finance need to be identified if such an objective is to be met.

Another component of fairness is *status quo fairness*--or a non-loser criterion. No stakeholder is to be adversely affected in any way to implement this concept of fairness. Of course, such an

objective could be extremely costly in terms of lost opportunities and reduced productivity growth. However, the idea of a "safety net" is relevant in some situations.

Cost causality is another component of fairness. From this perspective, those who cause costs to be incurred ought to bear the burden of paying those costs. Regulators may selectively choose to set certain prices below the cost of the service. This does not imply that the regulator's knowledge of the true cost is inessential. Indeed, no "social planner" can perform his duties effectively without full knowledge of the resultant costs and benefits. The key point is that knowledge of respective service costs is important, if only to know what it is that rates are departing from.

Finally *fairness in process and opportunity* is another dimension warranting attention. Open hearings and the elimination of uneconomic entry barriers are based on the principle that competition serves us well: in the marketplace of ideas and for most goods and services. That these four conceptions of fairness may conflict with one another at times is not surprising. We are continually balancing outcomes on the basis of how we value their various aspects. The key is to make those trade-offs explicit, so we understand what we are giving up when we choose one outcome over another.

3. Assurance of Opportunity to Obtain Revenue Requirements

The task of setting revenue requirements closely parallels that of rate design--with one key difference. Allowed returns and legitimate operating costs establish the level of revenue requirements while rate design is concerned with achieving efficiency, with how to obtain cash flows, and which class of customers bears the greatest "burdens."

The regulator is first and foremost required to set the revenue requirements for the firm so as to ensure its ability to raise capital. That is, the regulator must ensure the viability of the enterprise. Historically, regulators and the regulated have not always agreed on the capacity investment's size or how it should be financed. In general, competitive entry will tend to add an

additional constraint to the process. Not only will competition serve to highlight past departures from competitive market pricing, it will also ensure that such departures not continue.

The financial viability of firms is essential if investments needed to meet consumer demands are to continue. Both case law and common sense dictate that the providers of telecommunications services be given the opportunity to recover costs incurred to meet customer demands. Regulation does not "guarantee" investors a rate of return on their investment, but regulatory constraints must not preclude cost recovery of prudently incurred expenses.

There is another issue of potential importance, as telecommunications markets become increasingly competitive. This issue centers on the possible legal challenge to the regulator's discretion in setting both overall revenue requirements and rate levels necessary to generate these revenues. For example, a regulator's decision to simultaneously allow competitive entry into a given market while constraining the regulated firm to rigid pricing and service obligations could be viewed as an unjust confiscation of property in violation of 14th Amendment guarantees of equal protection.

4. Fairness in the Apportionment of Total Costs

The rate design function critically depends on understanding the underlying cost structure for the service in question--as every departure from incremental cost results in a loss in economic welfare relative to a competitive marketplace. From an economic efficiency perspective, unallocable costs should be borne by services that will minimize the divergence of demand patterns from those that would have been observed in a competitive marketplace.

As a general guideline, regulators may want to ensure that all services cover the cost to the supplying firm. Any rate that then covers the firm's cost of providing a service is understood as *compensatory*. Of course, as discussed above, regulators may ultimately decide that fairness issues warrant a departure from such a standard--but without such cost information they could not know what they were departing from. This is a particularly important consideration under rate of return

regulation, where the cost not paid by one class of customers must be passed on to some other class.

Note that although consumption pattern distortions may be minimized if inelastic demanders are priced to recover large proportions of unallocable costs, those demanders may be precisely the ones regulators most want to protect. Inelastic demanders have few substitutes for the service in question. Yet, as Henderson and Burns stress in their recent NRRI report, value-based pricing may be appropriate in the presence of shared (unallocable) costs.[9] Thus, the identification of "undue" price discrimination turns on alternatives available to "favored" customers in the long run. Henderson and Burns conclude that traditional (historical embedded) cost of service studies are inadequate to distinguish "due" from "undue" price discrimination (p. vii).

The advent of competition places an additional constraint on the regulator's ability to establish rate levels for services. In a competitive marketplace, inefficient rate levels are synonymous with unsustainable rate levels, which generally means that the firm will fail to attain the particular revenue requirement either in aggregate, or for particular service classes. For example, prices above stand-alone costs will tend to induce self-provision of a service. Similarly, prices above costs of serving a coalition of customers (e.g., urban high volume customers) will be unsustainable in the long run, as new firms or groups of users enter the industry and provide service offerings. Such defections may be economic (if lower-cost suppliers are involved and the incumbent does not lose multiproduct economies) or uneconomic (as with "cream-skimming" induced by mandated rate-averaging for a regulated incumbent).[10]

5. **Avoidance of Strategic Behavior**

Regulators may want to attempt an interesting experiment. Require each party to a regulatory hearing to submit a particular cost methodology. Each firm desiring business in the telecommunications market in question must blindly draw a cost methodology at random from those supplied and utilize it in decision-making. We might be amazed at how similar these cost

methodologies are and, perhaps more importantly, how closely they parallel principles applicable

to a competitive marketplace.

This last issue deals with the manner in which both competitors and market incumbents can be expected to use costing principles that support and reinforce their own strategic best interests. For example, the market incumbent may be expected to propose a price floor equal to incremental cost. Existing competitors and potential entrants may object to this. First, they might argue that incremental costs are such that entrants would be closed out of the market if the incumbent had such pricing flexibility. Second, competitors may argue that since the incumbent is "guaranteed" capital recovery, the incumbent will over-invest in order to drive marginal cost down to artificially low levels and thus preclude competitors from the marketplace.[11] This is one reason why potential competitors might argue for some type of fully allocated costing methodology for the incumbent--so as to preclude this type of strategic behavior.

The incumbent could argue that its costs are real and that with over-investing, the regulator is responsible to identify it and make appropriate disallowances. Hence, it is unreasonable for incumbents to be precluded from a sale that would have contributed to the recovery of fixed and overhead costs.[12] There is no straightforward solution to this conundrum. In fact, with increasing competition, it will become increasingly more difficult to discern over-investment from stand-by capacity necessary to meet carrier of last resort obligations.[13] It may be best to use some measure of long run incremental cost as the basis for a price floor and depart from that benchmark on a case-by-case basis.

Regulatory Tasks in a Changing Environment

The historical role of the regulator, at least in theory, was a substitute for the competitive marketplace, which was unreliable to produce acceptable performance under conditions of natural monopoly. To this end, regulators monitor the incumbent to ensure acceptable performance--defined

in terms of the regulator's public policy goals. This task has now grown increasingly more complex at the state level due to technological advances, divestiture, FCC policies, and resultant competitive entry. The situation may now require the regulator to serve the role of "central planner." However, the role of umpire and incumbent-handicapper is not easy, as the length and complexity of regulatory proceedings illustrate.

Cost information will play an increasing role in regulators monitoring market activities in multiple dimensions—only a few discussed above. Determining how this information is utilized and establishing the standards preserved by it will nonetheless be somewhat subjective. It is our view that in deciding to pursue a certain policy objective, regulators should at least be apprised of the applicable economic costs and benefits.

The Economic Standard: Marginal And Incremental Cost

The distinction between marginal and incremental cost is one of degree.[14] By marginal cost, we mean the change in the firm's total cost resulting from an infinitesimal change in output—generally one unit. With respect to some technologies, including those governing most regulated industries, it may not be meaningful to discuss one additional unit of output. In these circumstances, output (and capacity) tends to be added in indivisible quantities. Hence, when the additional output is more than one unit, we refer to the corresponding cost measure as average incremental cost, or simply incremental unit cost. By this we mean the firm's total costs (from supplying the increment in output) divided by the number of units that increment is capable of supplying. For example, an optical remote switching unit may be capable of switching 10,000 lines. In this circumstance, it makes little sense to refer to the marginal cost of a single additional access line. It makes more sense to look at the "lump" of investment in total, conditioned on the number of units this increment is capable of supplying. Hence, if this optical remote unit cost $2 million, then the long run incremental unit cost per access line is on the order of $200.

Economists generally support the use of incremental cost as a benchmark for pricing due to its efficiency properties.[15] Since consumers use prices to guide their purchasing decisions, every price deviation from marginal cost results in a welfare loss to society--meaning that voluntary transactions which benefit the firm and the consumer are discouraged. These principles are illustrated in Figure 1 for P > MC. Let "D" in the diagram represent a standard downward sloping demand curve. Quantity demand is greater at lower prices than at higher prices. Let "MC" represent marginal cost and P_0 represent the initial price level. The triangular area marked with an "L" represents the welfare loss from price being set in excess of marginal cost. Alternatively, this area represents the market valuation of those transactions that were discouraged because of inefficient pricing--sometimes referred to as *allocative inefficiency*. Note that these welfare losses--frequently referred to as the deadweight loss--is symmetric with respect to P > MC and P < MC.

[Figures 1 and 2 About Here]

Another source of inefficiency is known as technical (or production) inefficiency where society is not using the least cost method of production. This is illustrated in Figure 2. MC^R and MC^E, PR and PE represent the regulated firm's and the entrant's marginal costs and prices, respectively. When the regulated firm's price is pegged above marginal cost (perhaps in order to provide a contribution to covering some shared costs) competitive entry may occur. Such entry is inefficient for the costs shown here, i.e. the shaded area in the diagram represents the excess resource costs borne by society in producing output Q^E with the relatively inefficient production process. Since consumers make consumption decisions on the basis of price signals (rather than cost signals), they are led to substitute the lower price (higher cost) service for the higher price (lower cost) product. Note that the socially efficient output level is Q^*.

The Bell operating companies have argued that when their rates for selected services are set far above underlying incremental costs, uneconomic bypass results as lower price (higher cost) services
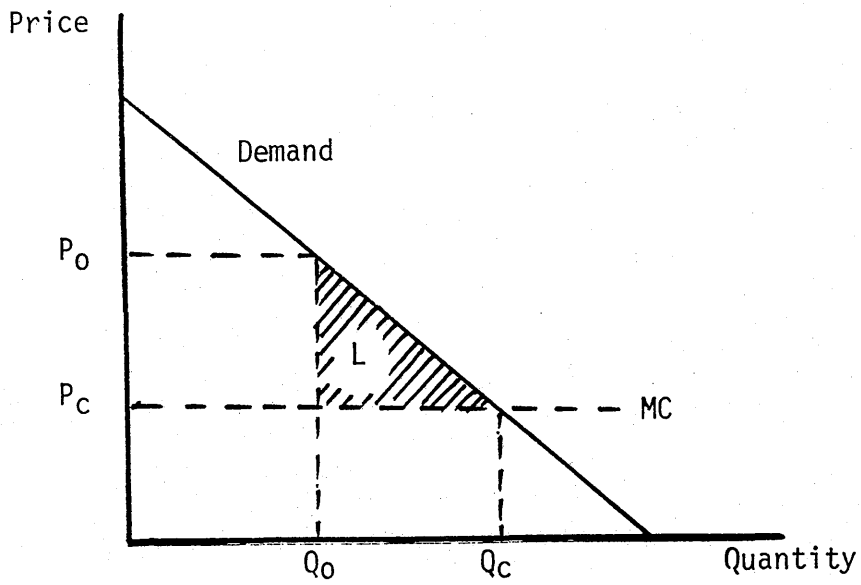
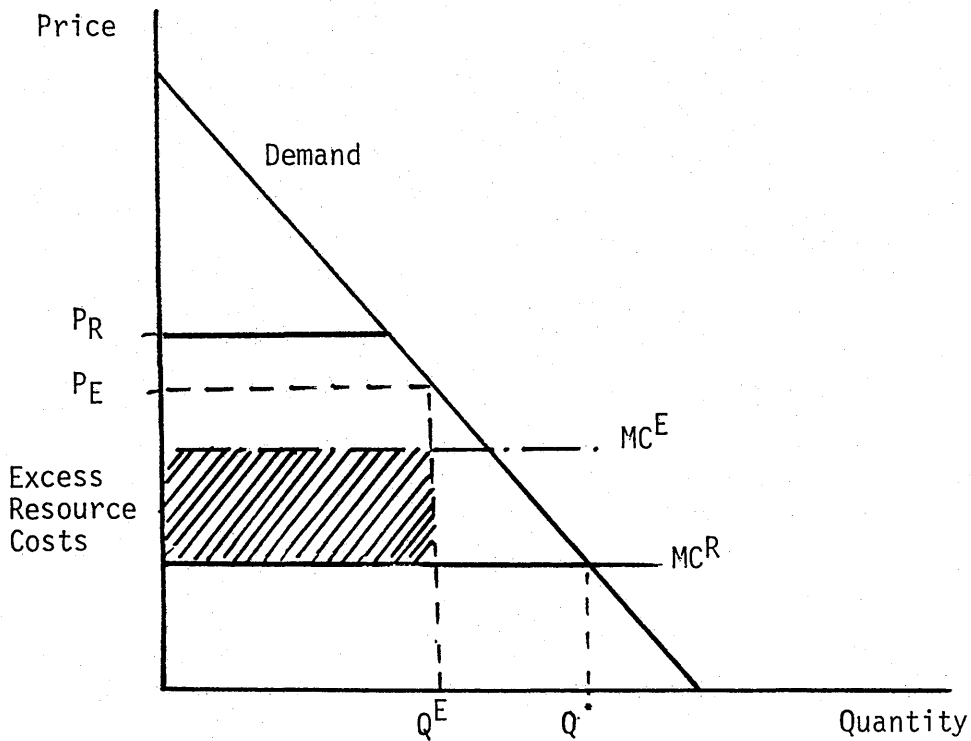FIGURE 1
Allocative Inefficiency

FIGURE 2
Technical Inefficiency

are substituted for higher price (lower cost) services. This is an example of *technical or production inefficiency*. The foregoing analysis reveals the premise underlying the economist's support for marginal cost pricing: it sends the proper signals to consumers, yielding efficient resource utilization. Departures of price from marginal cost are revealed to encourage either under or over consumption relative to what is efficient from society's perspective. In addition, P > MC can induce higher cost entry which produces output displacing efficiently-produced output.

Of course, regulators may wish to encourage competitive pressures which stimulate cost-cutting activity by the incumbent (regulated) firm. *Innovative efficiency* may be enhanced by potential entrants who are not committed to old production technologies. Regulators will want to encourage technological change which leads to meeting new needs or serving old demands at lower incremental costs. A comprehensive examination of this issue is beyond the scope of this study. Suffice it to note that regulatory policy may face a trade-off between technical efficiency (relative to today's technology and input prices) and innovative efficiency (discovering new services and introducing new production processes).

The incremental cost discussion is framed within the context of "cost-causality." Optimal resource allocation generally requires that buyers be responsible for the costs that they cause to be incurred. Conversely, if buyers of a new service are held responsible for costs less than that which they cause to be incurred, buyers of the firm's other services must be subsidizing the new service's production. In addition, over-consumption results as buyers of the new service are signaled that the resource cost of their purchases is less than cost. Consequently, prices set either above or below incremental cost induce different market behavior from prices set equal to incremental cost.

Cost causality can also be examined from the standpoint of an entire product line. If it costs $100 to produce 100 units of the regulated service, and $120 to produce those 100 units *and* 10 units of an additional service, then the *total service incremental cost* is $20, and the average service

incremental cost is $2. If the new service had a marginal cost of $1 (that is, the cost of increasing output from 9 to 10 units), pricing the service at $1 (and selling 10 units) would not be compensatory. Regulators would want to ensure that the new service recovered the total service incremental cost, $20. This would be accomplished via multipart pricing, where the marginal price is $1, but some fee or access charge is allowed to cover the total service incremental costs of the product line. As Baumol notes, the regulatory objective is to avoid burdening the original customers with costs incurred to serve buyers of the new service.[16] Yet, it is equally true that "new services cannot cause old costs."

Of course, the above situation is very simplistic--with no cost complementarities between the new service and the old one. That is, production costs of the old service are unaffected by the addition of a new service and by new service output changes. If production costs of the new service are lower when produced in conjunction with the first service (due to, say a shared input) than when produced separately, then fairness might dictate that the new service's consumers contribute to covering some costs of that shared input. However, efficiency does not require such a transfer (or cost allocation of "old" costs). Furthermore, an inefficient production configuration could be induced by too high a *tax* or contribution.

Another complication is the possibility that the new service is a substitute for, or a complement of, the initial service. If the new service is a substitute for the regulated service, a net-incremental revenue test is needed to assure that lost revenues (and avoided costs) associated with old service output reductions are taken into account. In addition, when many services and customer groups are involved, the test must be extended to combinations of services. While these cost and demand interdependencies complicate regulatory analysis, the principles are straight-forward. Fully distributed cost concepts *do not* provide a good indication of whether burdens are being imposed on customers of the initial service when new services are introduced.

Short Run and Long Run Incremental Cost

To this point in our discussion, we have used the term incremental cost as if the term were well-defined. In reality, the marginal cost concept has generated significant debate--not only in the hearing room, but in the academic literature as well. The debate continues, as reflected in the 1991 NRRI volume. While we will not resolve all these difficulties, we will shed light on some areas of broad agreement and disagreement.

A major source of debate and confusion with the marginal cost concept is the distinction between short run and long run incremental cost, and which one appropriately measures economic efficiency.[17] By long run incremental cost, we mean the change in the firm's total costs resulting from a one unit change in output, when all inputs are perfectly divisible and optimally employed. By short run incremental cost, we mean the change in the firm's total costs resulting from a one unit change in output, when the firm cannot vary all its inputs.[18]

Short run incremental cost has acquired one meaning in the hearing room and quite another in the academic literature. In the academic literature, it refers to the incremental change in the firm's total costs when there is a change in output, but some of the input quantities cannot be varied. Conversely, in regulatory hearings, it generally means the change in total costs when there is excess capacity for production.[19] In the former case, short run incremental cost exceeds long run incremental cost over particular ranges of output. Whereas, with excess capacity, short run marginal cost is exceeded by long run incremental cost.

For the moment, let us work examine the regulatory definition of short run incremental cost. To clarify some theoretical constructs while simultaneously recognizing some practical considerations, we consider a stylized example from the telecommunications industry. Suppose only three inputs are required to supply one unit of basic residential telephone service (1FR): *Switching, Feeder Plant and Distribution Plant*. Each 1FR requires two units of switching, and one unit each of feeder and

distribution plant. Suppose that input costs are as follows: switching, $3 per unit; distribution, $4 per unit; and feeder, $2 per unit. Long run incremental cost is thus $12.

Now suppose that switching capacity is perfectly divisible, and there is excess capacity for both distribution and feeder plant. Short run incremental cost is thus $6. Let us assume six prospective subscribers of 1FR service with respective valuations of $35, $16, $10, $8, $7 and $3. In scenario I, suppose that the 1FR price is initially set equal to long run incremental cost. In this case, consumers 1 and 2 subscribe, but consumers 3-6 do not. Total revenues equal $24. In scenario II, suppose that the 1FR price is set equal to short run incremental cost. In this case, all consumers subscribe, with the exception of consumer 6. Total revenues thus equal $30.

The question is whether "first-best" efficiency is obtained with prices set equal to long run marginal cost or to short run marginal cost? Alternatively, is there any economic efficiency justification for denying 1FR service to consumers 3-5? Let us examine both sides of the issue.

Long Run Incremental Cost

> *Long run incremental cost is the resource cost incurred by the firm in supplying an additional unit of the commodity when all inputs are perfectly divisible and optimally employed.*

Supporters of long run marginal cost as a benchmark do so on the grounds that efficiency is attained when prices reflect the corresponding resource cost to society for supplying the next unit of output.[20] For such analysts, the fact that the firm may be operating with excess capacity in some of its inputs (i.e. those input costs are sunk) is no reason to signal consumers that the resource cost is less than long run incremental cost. In fact, to do so would send the improper signals to consumers as they make long run investment decisions in durable goods based on artificially low prices.[21]

The long run/short run distinction may not be the most instructive for understanding operative subtleties. *Full adaptation* versus *partial adaptation*, may more meaningfully convey the actual situation. Full adaptation refers to an environment in which the firm is producing *de novo* with no

sunk inputs.[22] This concept parallels the long run incremental cost measure but is not inherently time related. Partial adaptation recognizes that the firm may have sunk inputs and face other constraints (eg. fixed plant size) that require an output production process not otherwise chosen if it had the ability to adapt perfectly to existing market conditions. A partial adaptation mode entails both production constraints (i.e. inability to use least-cost production techniques) and "free inputs" on the margin (i.e. sunk costs not included in incremental cost). Recognition of partial adaptation may generate incremental cost measures that are greater than or less than full adaptation incremental cost estimates.

It is argued further that when prices are set equal to short run incremental cost, they can be characterized by volatility that poses institutional difficulties for both consumers and producers. On these grounds, it may be prudent to set prices on the basis of long run marginal cost. However, Edward Park finds that variable prices (reflecting short run conditions) can offer substantial efficiency gains for some situations involving lumpy telecommunications investments.[23] The revenue requirement constraint results in the efficiency gains being passed on to consumers.

With respect to our 1FR example above, advocates of long run incremental cost would presumably argue that price should never be set below $12. Hence, according to this standard, only consumers 1 and 2 should be supplied with service. For capacity planning purposes, expected future consumption depends on the forecasted price--and a price reflecting only short run incremental costs could induce excessive investment in capacity to meet the demanded quantity. However, once the lumpy investment is made, those sunk costs can be ignored for purposes of signalling additional consumption at that time. As noted above, allowing variability in prices, reflecting short run considerations, can enhance efficiency.

Short Run Incremental Cost

> *Short run incremental cost represents the resource costs borne by the firm in supplying an additional unit of output when using the existing technology and taking into account that some input costs are sunk and thus not avoidable with respect to providing the next unit of output on the margin.*

Supporters of short run incremental costs do so on the basis of demand uncertainty and efficiency. Alfred Kahn, Lester Taylor and others emphasize the uncertainty of demand forecasts associated with long run capital investments in the telecommunications industry.[24] Hence, at any given point in time, capacity may exceed or fall short of the optimum given prevailing demand conditions. Consequently, from this perspective, prices should rise and fall to ration demand to existing capacity at any given point in time.[25] To insist upon using long run marginal cost under conditions of excess capacity blurs the distinction between sunk and avoidable costs with corresponding losses to economic efficiency.[26]

Denying 1FR service to consumers 3-5 is equivalent to assuming that the necessary feeder and distribution plant costs could be avoided if these consumers did not subscribe. Clearly, this is not the case since these costs are already sunk. Hence, if in a "first-best" world prices signal the resource costs to society of supplying the marginal unit of output, then feeder and distribution costs should not be included in this particular cost calculation. Yet, at the point in time the utility considers additional investment in the network, *all* consumers of the service in question should be levied a price that reflects the true resource or opportunity cost of continued consumption. Hence, the appropriate cost standard under such conditions is long run marginal cost. Why? It is only consumers who do not terminate consumption when confronted with a price greater than or equal to long run marginal costs that are *causally-responsible* for such capital costs.

Space limitations preclude resolving all these issues. Indeed, this debate has continued for well

over a century, and we are not likely to resolve it here. We confine our attention to a few observations.

First, it is important to understand the rationale for the choice of short run or long run incremental cost. For instance, we may agree that short run incremental cost is the proper benchmark for economic efficiency and yet simultaneously recognize that the volatility in the resulting price path is unacceptable for practical or institutional reasons. In other words, it is important first to agree on what measure sends the correct signal, and then depart from that measure as dictated by the needs of a practical and consistent regulatory policy.

Second, a critical interplay exists between the choice of marginal cost measure and the operative form of regulation. Should regulators allow the firm to price on the basis of short run incremental cost, the question remains, who pays for the feeder and distribution costs; while not part of marginal costs in a forward looking sense, they remain part of the firm's revenue requirement which must be paid by subscribers in the aggregate. This may be where the *stand-alone cost* concept bears some fruit. For example, while regulators may be sympathetic to allowing the firm to price down to short run incremental cost, that same regulator may be concerned that some other ratepayer, or class of ratepayers, must foot the bill for the difference between short run incremental cost and the average revenue requirement for that particular service. Hence, regulators may want to establish bounds serving both economic efficiency and fairness. This would involve granting the firm flexibility to price down to short run incremental cost, yet imposing the requirement that no class of consumers pay more than the stand-alone costs associated with their service. Multipart pricing is useful in such situations.

Third, as discussed in greater detail below, it is important to understand the strategic dimensions associated with unlimited pricing flexibility. Indeed, the combination of pure rate of return regulation and pricing flexibility down to short run incremental costs may generate some rather perverse

investment behavior that serves neither fairness nor economic efficiency. Efficient potential entrants with lower long run incremental costs for a new service could be prevented from entering the market if the incumbent were allowed to price a substitute service at short run marginal cost. This may dictate the need for some combination of *price caps* on existing services, and *price floors* on new services. In the process, regulators could explicitly rank services on the basis of fairness and distributional concerns.

Fourth, the comparison of long run incremental costs across firms establishes efficiency superiority. A comparison of short run incremental cost measures conveys no informational value regarding the relative long run efficiency of the firm. This issue concerns proper comparisons of technical inefficiency in production.

Finally, the operative competitive dimensions must be taken into account. While regulators must be concerned that the regulated firm not be granted flexibility that impedes otherwise socially beneficial competition, it must also ensure that competitors do not use the regulatory process to secure strategic advantage not otherwise attainable in the marketplace. In other words, competitors must not be allowed to arbitrage the regulatory process to obtain private benefits which are exceeded by resultant social costs.

Some Objections To the Use of Incremental Cost Pricing

Having discussed the efficiency properties of the incremental cost benchmark, some objections to this standard are reviewed.

Regulatory consent allowing the regulated firm to price down to short run incremental cost may entail a transfer: a shift of revenue requirements to another class of ratepayers. How might this situation be addressed?

The short run incremental cost standard for pricing flexibility raises some fairness and income distributional concerns. Specifically, as mentioned above, regulators may recognize the efficiency

properties of short run incremental cost, yet also realize that average revenue requirements must be paid to keep the firm viable. How are these issues resolved? We believe that there are some *efficiency gains from self-selection* in this particular case. Specifically, the regulator can provide the firm with the choice of pricing down to short run incremental cost, without guaranteeing that average revenue requirement burden will shift in total to some other class of ratepayers. Mechanically, this can be handled by capping the rates on some services, while simultaneously permitting the firm to choose its desirable strategy. For example, pricing down to short run incremental cost may result in under-recovery of the revenue requirement which, given caps on the prices of other services, needs to be recovered from shareowners. This is precisely what would happen in an unregulated market, and we believe it preserves some desirable efficiency properties as well.[27]

The incremental cost approach will generally fail what we might call the "adding-up" property. (We discuss this further in the next section.) Under standard approaches it will be difficult to work back from incremental cost and account for all costs incurred by the firm. However, this would not appear to be an insurmountable problem. The firm can track its costs so that regulators can add-up to totals without such totals directly affect pricing decisions. In this regard, it is important to recognize that one costing system or method is unlikely to provide regulators with all information needed to prudently regulate and monitor the firm. Historically, the problem has not been the existence of numerous costing methodologies, but the recognition that these diverse methodologies are not interchangeable. That is, cost accounting reports that track the expense of a given product line may be quite beneficial for regulators in understanding the level (and mix) of expenditures. Expenditures are placed in accounts that can provide information regarding directly attributable costs. It would be a mistake, however, to presume that this same costing methodology can play a role in efficient rate design, particularly if arbitrary cost allocation procedures are involved.

Objections to the long run incremental cost measure will be lodged by competitors who claim

that such a standard is anti-competitive. The history of the railroad industry, particularly in competition with trucking, is filled with attempts by competing entities to saddle each other with burdensome cost standards for rate design. The same behavior has posed, and will continue to pose, a dilemma for regulators in the telecommunications industry. Several observations are in order.

We noted above that a comparison of short run incremental cost measures is of no informational value in assessing relative efficiency across firms. It is necessary to analyze long run incremental cost measures for this purpose. Hence, as a benchmark, regulators may want to adopt a standard of long run incremental cost and require the regulated firm to demonstrate why deviations may be socially beneficial. If regulators believe that a short run incremental cost standard will impede otherwise socially beneficial competition, then this approach may be warranted. In the longer run, if such competition is not welfare-enhancing, then the marketplace will sort this accordingly. Moreover, we again point out that no party should be allowed to secure in the hearing room what is unattainable in the marketplace. Arguments that the regulated firm should be subject to a strict fully distributed cost standard are totally without merit. No efficiency or equity ends are served by such a standard, and the welfare losses are likely to be extreme. In addition, claims by competitors that costs are being off-loaded to other, perhaps less competitive services, is a frequent refrain. There is no question that under rate of return regulation, the firm may have selective incentives to engage in such behavior.[28] Nonetheless, such incentives are largely muted by adopting appropriate rate caps on other services or groups of services. This procedure will largely eliminate any incentive for the regulated firm to engage in such behavior, as there is no longer any return from shifting costs to services whose rates cannot be altered to recover them.

Moreover, a number of authors object to the use of marginal cost pricing because it may be subject to strategic manipulation. Wilson (John W. Wilson, "The Strategic Value of Marginal-Cost Pricing," in Pollard, ed., op. cit., pp. 545-557.) contends that the *theory* of marginal cost pricing is well-

defined, but problems arise when it comes to actual *practice*. Specifically, there can be serious questions as to which service is truly marginal. In addition, Albery and Sievers (Mark Sievers and Brooks Albery, "Using Incremental Costs to Detect or Prevent Predation in the Long-Distance Telecommunications Industry," in Pollard, ed., op. cit., pp. 619-647) contend that dominant firms may engage in so-called strategic allocation of overhead and thereby conceivably drive a relatively more efficient, albeit non-dominant, competitor from the marketplace. The upshot is that regulators must go beyond simple enforcement of a marginal cost standard if the long term objective is dynamic efficiency and maximal economic welfare.

Actual measurement of marginal cost is complex and has been the subject of considerable academic debate and regulatory discussion.[29] We only note that getting the measure 90 to 100 percent correct is difficult. An 85 to 90 percent correct measure is probably attainable for most telecommunications services, and the incremental efficiency gains from the extra 10 to 15 percent accuracy improvement are likely to be sacrificed in the form of rent-seeking within the context of academic and technical debates. It is not difficult to be approximately right: we believe this to be the appropriate target. It is also not difficult to obtain precise calculations of wrong numbers. Such calculations (as with embedded cost studies) may help regulators evaluate developments retrospectively, and they may be necessary to obtain overall revenue requirements. However, fully distributed cost studies based on historical costs provide no guidance for efficient pricing.

Cost Allocations and the Regulatory Process

Cost allocations have a long and venerable history, in the telecommunications industry and in all regulated industries. However, cost allocation systems are not confined to industries subject to governmental oversight. Many businesses in competitive markets use cost allocations systems as well. Before delving into the underlying cause of using of cost allocations, let us briefly describe what they are and how they work.

By a cost allocation system, we generically mean any costing methodology that apportions costs incurred by the firm to various activities and products such that: (1) all costs are accounted for; (2) costs are distributed across products and services with little or no regard for cost-causation; and (3) the "adding-up property" is always satisfied. The critical point is that cost allocation systems are first and foremost an accounting tool, and as an accounting tool they may serve a perfectly useful and even socially beneficial function. When these systems are used for other purposes, such as establishing prices, efficiency problems arise.

Economists have long been opposed to cost allocation systems--largely because such systems do not distinguish between *sunk* and *avoidable* costs. Consequently, inefficiency is perpetuated because prices based on fully allocated costs are incapable of signaling the true resource costs to society associated with the product or service in question. This is not to suggest that cost allocation systems have no redeeming qualities--in fact, they do, but economic efficiency is not one of them.

Consider the following debate concerning additional access lines, or so-called second lines. Some telephone companies have contended that the long run incremental cost of basic local service (1FR) is on the order of about $25 per month. Nonetheless, telephone companies have begun to selectively market additional lines (that is second lines into a residence) at prices of between $12 and $14.

The telephone companies contend that in selected areas, excess capacity exists in feeder, distribution and switching. Hence, the true incremental costs of supplying the second line should not treat such costs as avoidable. By selling additional lines at prices exceeding the true incremental costs, yet discounted from the standard tariff rate, some contribution is made to the recovery of fixed and overhead costs. The telephone company and its ratepayers are thus better off as a result of these actions.

If cost allocation systems are so problematic, why do so many firms use them? We offer a

few thoughts on this matter. First, the role of costing in most firms has traditionally been an accounting function, and the number one task of most accountants is to ensure that all costs are accounted for. On this score, cost allocation systems receive high marks. They allow managers to track every dollar of costs and thus satisfy the so-called "adding-up" property. Second, when demand is tracking expectations, input prices are stable, and technologies are unchanged, no deviation may exist between what fully allocated costing systems and what incremental cost systems propose for prices. It is precisely when competition makes inroads into a market, or when economic change occurs, that the recommendations from these two approaches differ, and then in substantive and significant ways. For example, in the face of serious downturns in demand, either because of competition or some exogenous factor, output or revenue-based cost allocations systems can serve to exacerbate the problem of covering costs.

Concluding Remarks

In this survey of costing principles, we have attempted to provide the reader with an overview of the costing issues currently under debate in the telecommunications industry. Since cost data generally serves as the foundation for rate design, these questions will surface in regulatory hearings. Competition will serve to fuel what is already a highly litigious process--and much of this hearing room debate will focus on the appropriate principles for cost measurement.

Some issues we discussed, such as the efficiency properties of incremental vs. fully distributed costs, will be accepted without exception by most economists of academic repute. Other issues, such as the long run or short run marginal cost standard and sufficient costing information for protection from cross-subsidization, are more elusive and require more technical analysis than presented here.[30] These issues warrant thoughtful, independent study.

Finally, any study of costing issues recognizes at the outset that economic analysis often produces a grey answer when regulators desire black or white. Such knife-edge precision is probably

inappropriate for cost analysis when production technologies utilize shared inputs--as cost responsibility is difficult to assess. Fully distributed cost methods enjoy some support because of the false sense of security that comes from (misleadingly) placing costs in one bucket or another. In other words, we may not have all the costs in the right buckets, but at least we know where they are. The grey areas in cost recovery analysis--meaning recovery of costs that cannot be directly attributed-- can be addressed through techniques of Ramsey, and multi-part pricing. In other words, it is essential for economic efficiency that the grey areas in cost analysis remain grey--in spite of the discomfort policy makers may feel with such an outcome.[31] In our view, a move in the direction of fully distributed costing methods is a move away from economic efficiency.

In summary, we note that the recent *Report of the Blue Ribbon Telecommunications Task Force to the Illinois Commerce Commission* includes many points which parallel those presented here. The Task Force recommends that the Commission be authorized to justify prices by a variety of standards, including new service prices that cover long run marginal cost. As the *Report* states:

> ... confronted with the need to set rates for services provided with common plant, regulators have traditionally justified prices on the basis of some allocation of common costs. Apparently, bowing to what they perceive as the demand of legislators, judges, and laymen to cost justify prices, regulators have pretended what must be done arbitrarily can be done in a principled fashion. They allocate common costs among services and among time periods. They pretend to estimate a separable cost for each service even though no such cost exists. Regulators have employed a variety of allocation methods. Some appeal to different notions of "fairness" more than others, but all rely on fictions. These fictions tend to obscure what regulators are really doing. Moreover, their use fosters the notion that "cross-subsidy" has the same unambiguous meaning in the presence of common costs that it does in their absence. It does not (except tautologically), and fostering the notion that it does makes for mischief.[32]

The present study is consistent with the thrust of the Illinois Task Force Report, and with its emphasis on incremental costs as the basis for evaluating efficiency. When conducting a costing study, it is better to have a rough approximation of the right number than a precise calculation of an irrelevant number. Estimates of incremental costs are the relevant starting points for efficient pricing in telecommunications.

Footnotes

1. Meitzen, Mark E., "Diversification of Telephone Company Service Offerings and Cash Cow Economics," *Utilities Policy*, Vol. 1, No. 1, October 1990, 43-53.

2. In fact, the increasing number of attorneys, economists and engineers appearing in regulatory proceedings could well be interpreted as rent-seeking behavior, in which interested parties battle over the slices of a fixed pie—a zero or negative sum game. What is unclear is whether the quality of the regulatory outcomes is improved (or the size of the pie increased!) as a result of the abundance of technical experts appearing in support of a given cause. See Jack Wenders, *The Economics of Telecommunications*. Ballinger: Cambridge 1987; and Fred S. McChesney, "Rent Extraction and Rent Creation in the Economic Theory of Regulation," *Journal of Legal Studies*, Vol. XVI (January 1987).

3. Thomas McCraw, *Prophets of Regulation*, Belknap Press: Cambridge 1984, pp. 256-259. See also, Alfred E. Kahn, *The Economics of Regulation: Principles and Institutions*. Cambridge: The MIT Press, 1988.

4. Jack Wenders, op. cit.

5. Gerald Brock, *The Telecommunications Industry*. Harvard: Cambridge 1981.

6. William Pollard, "Overview and Insights from the Symposia," in Pollard, ed., op. cit. p. xi. A listing of titles and authors of the volume is a Who's Who of telecommunications research.

7. D. F. Greer, *Business, Government and Society*. New York: MacMillan Publishing, 1987, categorizes the problems justifying government intervention as involving market imperfections, market failures, dynamic incapacities, and ethical concerns. Imperfections are reflected in market power or information inadequacies: the conditions under which buyers and sellers come together deviate from competitive assumptions. Failures refer to externalities, public goods, and common property resource problems (eg. radio spectrum)—all of which characterize some telecommunications markets. Dynamic incapacities relate to the inability of institutional arrangements to adjust to new situations without causing undue hardship—microeconomic transition costs or macroeconomic instabilities. Finally, ethical concerns often motivate public policy: with the goal being fair (or less unequal) income distribution, the production of merit goods (such as education), and the achievement of other social goals.

8. Agreement of just what is in the public interest may be more apparent than real, particularly with regard to the ability of politicians to balance different objectives—such as fairness and efficiency. Our awareness of government shortcomings has increased over time. These limitations include government imperfections, failures, dynamic problems (myopia and delay), and inconsistencies in the application of ethical criteria. Imperfections refer to government as a monopoly, the existence of voter apathy, and persistent lack of information needed by officials. These problems, identified by Greer, op. cit., parallel imperfections observed in markets. Government failures arise when benefit-cost analysis is misapplied—primarily due to special interest effects, externalities, bundling of policies, and a tendency to impose uniform treatment despite diverse situations. Myopia and delay stem from intertemporal inconsistencies, when costs and benefits of policies differ in the long and the short run. Some analysts find that policy cycles follow the election returns: when in doubt, avoid inaction.

Finally, the stated goal of fairness is often used to justify regulation, although the actual incidence of the benefits and costs generally differs from popular understanding. The two alternative institutional structures for seeing that consumers' needs are met via least-cost suppliers (and that new goods and production processes are introduced) involve markets or governments: we either use markets to allocate resources or government agencies to allocate markets. Each institutional arrangement has strengths and limitations. The selection of the appropriate institutional structure involves a realistic comparison of performance with market incentives and with regulation.

9. J. Stephen Henderson and Robert E. Burns, *An Economic and Legal Analysis of Undue Price Discrimination*, National Regulatory Research Institute, NRRI, 89-12, August 1989.

10. Recent work by Gerald Faulhaber, "Cross-Subsidization: Pricing in Public Enterprise," *American Economic Review*, 1975, pp. 966-977, suggests that a price vector may not exist that allows the regulated firm to breakeven and yet is sustainable against competitive entry. To see why this result is so important, consider the following citation from Kahn, *The Economics of Regulation*, Wiley, 1971, p. 2231: "If a natural monopoly is producing and pricing as efficiently as possible, there is no need to bar competitive entry: It is economically unnecessary and will not take place anyhow." This was the traditional view of the problem prior to Faulhaber's analysis.

11. Consistent over-optimism by regulated firms with respect to future demand can be addressed by regulators in later hearings. In other words, it is essential that these cost changes be measured relative to some baseline scenario--fully accounting for the degree to which changes in demand and/or the product line caused capital investments to be either accelerated or deferred.

12. This is essentially the argument that AT&T used in its tariff 12 and 15 proceedings with the FCC.

13. Dennis L. Weisman, "Optimal Re-Contracting, Market Risk and the Regulated Firm in Competitive Transition," *Research In Law And Economics*, 12, 1989, pp. 153-172.

14. Bridger M. Mitchell, "Incremental Costs of Telephone Access and Local Use," in Pollard, ed., op. cit., pp. 191-309.

15. Lewis J. Perl and Jonathan Falk, "The Use of Econometric Analysis in Estimating Marginal Cost," in Pollard, ed., op. cit., pp. 443-464.

16. William J. Baumol "Minimum and Maximum Pricing Principles for Residual Regulation," in A. L. Danielson and D. R. Kamerschen (eds.), *Current Issues in Public Utility Economics*, Lexington, MA: Lexington, Books, 1983, pp. 177-196.

17. Lester D. Taylor, "Calculating Short-Run Avoidable Costs as a Floor to Price in the Telephone Industry," in Pollard, ed., op. cit., pp. 465-498; and Nina W. Cornell, "Optimal Costing and Pricing Methodologies for Regulated Monopoly Telephone Companies," in Pollard, ed., op. cit., pp. 499-544.

18. The operative principle is summed up quite nicely by William Vickrey: "the marginal cost that is relevant to a pricing decision is a marginal cost of the output that will be affected by the pricing decision over the period for which that decision is to be considered not subject to possible revision. To attempt to import into a pricing decision considerations of fixed costs that will not be affected even indirectly by that decision is to chase a very wild goose indeed." William Vickrey, "The Fallacy

of Using Long Run Marginal Cost for Peak-Load Pricing," *Quarterly Journal of Economics*, June 1985, pp. 1333-1334.

19. William J. Baumol and Alfred G. Walton, "Full Costing, Competition and Regulatory Practice," *The Yale Law Journal*, Vol. 82, No. 4, March 1973, pp. 639-655.

20. See, Ralph Turvey, "Marginal Cost," *The Economic Journal*, June 1969, pp. 282-299 and M. Boiteux, "Peak-Load Pricing," Chapter 4 in *Marginal Cost Pricing in Practice*, ed. by James R. Nelson, Engelwood Cliffs: Prentice-Hall 1964. Reprinted in *Journal of Business*, Vol. XXXIII, No. 2, 1960, pp. 157-179.

21. This problem presumably could be addressed through explicit contracting between consumers and producers. If consumers are willing to make long run investments in durable goods, it would seem they would also be willing to make a long term commitment to use the utility's services. Hence, any inefficiencies and attendant risks stemming from prospective volatility in the price path could be eliminated through explicit contracting between the parties. Such contracts would simultaneously reduce demand uncertainty for the firm and price volatility for consumers.

22. While we generally acknowledge that in the long run there are no fixed costs, this run maybe so long as to be of no practical relevance. For example, with fiber optics, once the firm buries the cable, those costs are essentially sunk for the foreseeable horizon. In this case costs are fixed even in the long run--the planning horizon for the firm--and incremental cost estimates would acknowledge them as such.

23. Rolla Edward Park, "Incremental Costs and Efficient Prices with Lumpy Capacity: The Single Product Case," *Rand Corporation*, R-3723-ICTF, June 1989.

24. Kahn, op. cit.; Taylor, op. cit.; and Roland Andersson and Mats Bohman, "Short- and Long-Run Marginal Cost Pricing, on Their Alleged Equivalence," *Energy Economics*, October 1985, pp. 279-288. The importance of the nature of the contract between consumers and producers in determining the relevant marginal cost standard is discussed in Dennis Weisman, "A Note on First-Best Marginal Cost Measure: In Public Enterprise," *Energy Economics*, October 1991 (Forthcoming).

25. See Harold Hotelling, "The General Welfare in Relation to Problems of Taxation and of Railway and Utility Rates," *Econometrica* 6, No. 3, July 1938, pp. 242-269, on the rationing role of prices. Park, op. cit., provides simulation results supporting the use of variable (short run cost-based) pricing: "The largest welfare gains under either optimal constant or optimal variable prices occur when the initial capacity is about the same as the initial demand, and the lump size is large. Then charging a high price can produce large welfare gains by postponing the time when an additional lump of capacity must be installed." (Park, op. cit., p. ix.)

26. Recognize, however, that the short run marginal cost measure should include *usage*-sensitive, though not time-sensitive, depreciation. To the extent that current consumption, via usage depreciation, causes the firm to accelerate the point in time that it initiates additional network investment, the short run marginal cost measure should include the time value of money over the period for which the investment is accelerated. It is intuitive, though incorrect, to reason that since current consumption accelerates the exhaustion date of the existing network investment, the short run marginal cost measure should include the time value of money over the period for which the

investment is accelerated. This is incorrect because consumers with a willingness-to-pay greater than short run marginal cost, but less than long run marginal cost, would terminate consumption when confronted with a price equal to long run marginal cost. The demand of such consumers cannot be causally-responsible for accelerating the exhaustion date of the existing investment, since that consumption would be cut back with the higher future price.

27. Note the parallels here with Baumol's theory of "quasi-permanence of rate reductions" as a solution to predatory pricing concerns. William J. Baumol, "Minimum and Maximum Pricing Principles for Residual Regulation," in A. L. Danielson and D. R. Kamerschen , op. cit., pp. 177-196.

28. This would require, however, that the firm somehow be able to recover costs in one category unrecoverable in another, that the requisite financial incentives for such cost shifting exist, and that regulatory monitoring capabilities are inadequate to identify cost-shifting behavior. Richard Rozek, "The Over-Capitalization Effect with Diversification and Cross Subsidization," *Economics Letters*, 16, 1984, pp. 159-163; Martin Grace, "Shared Inputs, Over Capitalization, and Regulation," *Economics Letters*, 22, 1986, pp. 381-384.

29. See the discussion in the NRRI Volume regarding engineering process and econometric cost models.

30. See Sanford V. Berg and Dennis L. Weisman, "Cross-Subsidization and Price Predation in Public Enterprise: The Case of Telecommunications," Public Utility Research Center, University of Florida, Working Paper #179, June 26, 1991.

31. "The cost study made me do it," is considerably more palatable politically than "My judgement is that the burden should be recovered proportionately more from service X than service Y."

32. Calvin Monson, *Report of the Blue Ribbon Telecommunications Task Force to the Illinois Commerce Commission*, December 1990, p. 33.