COST CONCEPTS FOR UTILITY REGULATORS

Mark A. Jamison, Ph.D.
Public Utility Research Center
University of Florida

October 19, 2006
Introduction

Over the past century, the regulation of utility prices has largely focused on related prices to costs because of several factors. Early attempts to regulate prices without regard to costs resulted in prices being either so high as to make the regulation of no real consequence or in prices being so low as to make the utility enterprise commercially unviable. As Glaeser (1927, p. 108) explained based on his research of the situation, “The primary purpose of governmental regulations is to secure the balance of cost and income.” There was also a concern in some instances that absent cost-based prices, utilities could discriminate against politically or economically weaker customers to the advantage of customers whom the utilities wanted to favor (Bonbright, Danielsen, and Kamerschen, 1988, pp. 33-36).

Regulators are also concerned with issues of fairness and efficiency: Regulators will sometimes focus on how the utility company’s costs are distributed across services and will use that information to make a case that the distribution is fair and reasonable (Bonbright, Danielsen, and Kamerschen, 1988, pp. 482-483). At other times, regulators may be more concerned with economic efficiency, considering how customers and operators respond to the economic incentives that prices provide and so will seek to use cost information as a guide for developing prices that encourage efficient decision making (Kahn, 1988, I pp. 65-57). In other instances, the regulator may be seeking to subsidize a particular type of customer such as the poor and seeks cost information to help determine how much of a subsidy is needed and where the subsidy may be obtained (Bonbright, Danielsen, and Kamerschen, 1988, pp. 164-178).

More recently, when competition began to reemerge in utility markets, there was a concern that the utility might cross-subsidize its potentially competitive markets with profits from non-competitive markets, might engage in predatory pricing, or might favor its own services over rivals’ services with respect to access to essential facilities (Baumol, 1979; Trebing, 1984). Finally, regulators are concerned that accurate cost information is needed to establish prices such as for network interconnection, to ensure that competitive markets function efficiently (Um et al., 2004).

Some of these purposes for cost information may be at odds with one another. For example, there may be many views of what constitutes fair and reasonable prices, and some of these views may be in conflict with prices that are economically efficient. Furthermore, subsidy-free prices and fair prices may not be the same thing. As a result of these potential conflicts between purposes, regulators often find themselves balancing objectives by either explicitly or implicitly weighing their goals, for example, giving up a certain degree of fairness in order to achieve a greater amount of efficiency.

In addition, no single measure of cost can inform regulators about these tradeoffs. For example, some observers may not view economically efficient prices as necessarily being fair and reasonable. As we explain in more detail below, economic efficiency generally requires that prices reflect the services’ marginal costs. But concepts of fairness are sometimes based upon how evenly the historical costs incurred to provide services are apportioned across classes of service. Also, comparing prices with marginal costs does not fully inform us whether a cross-subsidy exists. This need to examine trade-offs in pricing objectives, coupled with the fact that
different objectives are associated with different measures of cost, means that regulators often need more than one measure of service cost and also need to know how to apply the cost information that they have.

The increase in the number of regulatory agencies around the world has resulted in a rise in the use of cost studies, with the predictable result that there are conflicting studies and disagreements over what are best practices. The purpose of this paper is to inform regulators in the development and use of service cost information, relying upon well-established principles and practices in the accounting and economics literature. More specifically we describe and define cost concepts typically found in utility regulation, with special emphasis on how the terms are used in regulating utilities. Wherever possible we offer definitions and descriptions both from the academic literature and from common usage. Furthermore, we make recommendations on how terms should be defined to avoid the confusion that can result when an expression is used multiple ways, or when multiple terms might be used to say the same thing.

The remainder of this paper is organized as follows. In the next section, we describe the two basic approaches to measuring costs, namely the accounting approach and the economic approach. We then describe the accounting approach. Next we address the economic approach by first describing the cost structure of the firm from an economic perspective. We then describe how the economic costs of the firm are measured. We conclude with summaries of how regulators have applied the economic approach.

**Basic Approaches to Measuring Costs**

There are two basic approaches to measuring costs for a utility company; the accounting approach and the economic approach (Jamison, 1998, p. 118). These approaches are different in several ways. The *accounting approach* encompasses the methods that rely upon the company’s accounting data as its basic source of information. These data represent costs that have been incurred in the past by the utility (Crum and Goldberg, 1998, p. 47). For example, the values found in an account for cable and wire facilities represent the dollars1 spent on assets of that type that the utility has purchased over the past several years and that have not been fully depreciated. In regulation, the accounting approach is primarily concerned with the assignment of costs to services in ways that seem fair and reasonable to the analyst or other observers (Bolter, 1978).

The accounting approach’s key source of information, an *account*, is an item on a utility’s financial statements such as administrative expense or international long distance revenues (Horngren, Sundem, and Stratton, 2005, p. G1). There are three basic financial statements that are of interest in regulation. The first is the *cash flow* or *flow of funds statement*, which records all of the cash inflows and outflows that result from the normal operations and projects that the operator undertakes (Crum and Goldberg, 1998, p. 57). Cash flows are of interest to regulators in part because some regulators use projected cash flows to establish *X*-factors for price cap regulation.2 Revenue and operating expenses related to projects and normal operations are

---

1 For convenience and unless otherwise noted, we express monetary values in “dollars.”
2 For an explanation of price cap regulation and *X*-factors, see King (1998), Sappington and Weisman (1996), and Bernstein and Sappington (2000).
recorded on the income or profit-loss statement, along with interest, taxes, and depreciation expenses (Crum and Goldberg, 1998, p. 52). Operating expenses are costs incurred for inputs that are used up within a year’s time. Assets (facilities, other property and investments, current assets, and deferred debts) and liabilities (stock, long-term debt, non-current liabilities, current and accrued liabilities, and deferred credits) are recorded on the balance sheet (Crum and Goldberg, 1998, p. 47).

The values in an account may represent historical costs or current costs. The values are called historical costs when they represent the dollars originally spent by the utility (Crum and Goldberg, 1998, p. 49). For example, if a utility once spent $100,000 for a particular coaxial cable, that cable’s historical cost would be the $100,000 minus its accumulated depreciation. Depreciation is the annual accounting charge for wear, tear, and obsolescence of assets (Horngren, Sundem, and Stratton, 2005, p. G3). In regulation, depreciation is also viewed as capital recovery, which is the spreading of an investment in assets over time to be recovered through prices (Kahn, 1988, I p. 105). Embedded or original costs are other terms for historical costs. Historical cost measures are generally found in the company’s accounting records and generally form the foundation for the accounting approach to costing. The current cost approach generally begins with the historical costs and applies inflation factors to express the historical costs in current monetary units (Crum and Goldberg, 1998, pp. 65-72).

In contrast with the accounting approach, the economic approach focuses on how changing the output of the utility will affect the total cost of the company (Bonbright, Danielsen, and Kamerschen, 1988, p. 393; Kahn, 1988, I pp. 63-66). This cost is the present value of the opportunity costs of the firm, for a specific future production period, and for a specific level of production (Alchian, 1959; Salinger, 1998). Because the economic approach is forward looking, it does not rely upon a particular database of cost data. Economic cost studies in telecommunications regulation traditionally relied upon projections of incremental cash flows associated with changes in output (Foster and Bowman, 1988) or econometric estimates (Evans and Heckman, 1983; Elixmann, 1988). In the incremental cash flow approach, the cost analyst would project the company’s cash flows under a base case scenario and under a scenario that includes the change in output. The net present value of the difference between these two cash flows represents the economic cost of the output change in this approach (Salinger, 1998). The econometric approach relies upon historical data to make statistical estimates of how variations in output affect total cost. More recently proxy cost methods have been developed to estimate economic costs (Jamison, 2002; Gasmi, et al., 2002, p. 10). These methods use computer models of how a company might engineer its network to estimate how changes in output affect the costs of the network.

A further difference between the accounting approach and the economic approach is how the approaches consider the cost of capital, which represents the returns that the utility must provide to debt holders and equity holders to ensure that they continue to provide capital (Kahn, 1988, I pp. 42-45). The economic approach always includes the cost of capital as a cost. The accounting approach will do so only if the analyst is estimating revenue requirement, which is the amount of revenue that the utility needs to receive from customers in order to cover its accounting costs and its cost of capital (Kahn, 1988, pp. 25-26). In accounting, the return to the equity holders is called profit or accounting profit (Horngren, Sundem, and Stratton, 2005, pp. 56-63). In
economics, the term profit is reserved for equity returns that are over and above the cost of capital. In the economic context economic rent is another word for profit.

Lastly, the economic approach generally includes opportunity costs, which are the alternatives that must be foregone by the utility if it provides the output whose costs are being measured (Kahn, 1988, p. 66). Sometimes opportunity costs are adequately measured by the projected cash flows and cost of capital, in which case the incremental cash flow approach and the proxy approach, if done properly, would estimate opportunity costs. In some instances, a decision to provide output causes the utility to forego strategic opportunities such as the opportunity to delay production decisions until more is known about the future. This raises the issue of real options, which might be explicitly included in an economic cost study in which the costing approach does not capture all of the firm’s opportunity costs (Trigeorgis, 1999).

The differences between accounting cost studies and economic cost studies are not simply differences in how they treat common costs, historical costs, and technical efficiency, as some have claimed. Rather, as we described above and describe in more detail below, the approaches are different in their purposes, data sources, and underlying assumptions. An exception was the set of costing rules established by the Canadian Radio-television and Telecommunications Commission (CRTC) in the 1980s (Bigham and Wall, 1989). The CRTC required Bell Canada to track ex post whether its estimates of economic costs proved true when the company provided services and recorded outcomes in its accounting records. This requirement led the company to adopt an approach to economic cost studies that was based on business and operating forecasts and that could be expressed in forecasted accounting data. The accounting cost studies were then, in effect, expressions of how the economic forecasts had played out in reality and not simply what a cost analyst might view as reasonable allocations of accounting costs. We are unaware of anyone else using these approaches to economic costs and accounting costs.

**The Accounting Approach to Measuring Costs**

As we indicate above, the accounting approach is primarily concerned with the assignment of costs to services in ways that seem fair and reasonable to the analyst or other observers (Bolter, 1978). This typically involves two basic steps, but the second step may be omitted for some forms of cost studies.

*Direct cost analysis*

The first step is to identify and assign *direct costs* or *directly attributable costs*, which are the accounting costs for production inputs that are only needed to provide a specific service or set of services (Horngren, Sundem, and Stratton, 2005, p. G3). The use of these terms is generally

---

3 *Technical efficiency* refers to whether a firm is operating at minimum economic cost for its level of output and quality of product (Leibenstein, 1966). Technical efficiency is also called *productive efficiency* or *x-efficiency*.
4 See for example, TERA (2006).
5 47 CFR Ch. 1 64.901(b)(2) 10-1-05 Edition.
6 These terms are primarily used in the context of accounting costs, but some people use this term when referring to economic costs. In the context of economic costs, direct cost is synonymous with service incremental cost.
limited to inputs that have their own identity for accounting purposes (i.e., their own account or subaccount). An example of a direct cost would be the hourly wages paid to a live telephone operator whose only job is to assist callers. The wages would be directly attributable to a service called operators assistance. Assignment refers to the act of designating an accounting cost amount as a direct cost of a particular service. An accounting cost study that stops at this first step is called a direct cost study or an embedded direct cost analysis if historical costs are used (Jamison and Brevitz, 1987).

The accounting concept of direct costs should not be confused with the economic concept of incremental cost, even though the two have much in common. Direct cost takes the amounts in an accounting record as given and, after determining that the cost is appropriate for regulatory purposes, simply asks which service or services use the cost items represented in the account. An economic analysis would ask whether other services might have contributed to the costs of these items. For example, a telephone company might upgrade its cable facilities to provide broadband services. The additional cost of upgrading the facilities would be an incremental cost of broadband even though some of the facilities might be used only to provide voice services. An accounting cost study would consider the accounting cost for the voice-only facilities to be direct costs of voice even though a portion of the costs were caused by the company’s decision to provide broadband (Jamison and Brevitz, 1987).

**Fully distributed cost**

The second step in the accounting approach is to allocate the remaining accounting costs across services. This typically occurs in two parts. In the first part, the analyst identifies accounting costs that appear to be reasonably attributable to particular services or to activities whose costs have already been directly assigned. For example, specialized computing equipment used primarily to manage telecommunications traffic is reasonably attributable to usage-based services such as international long distance and local calling, but it may be hard to directly assign the accounting costs to any one of these services. Such costs are sometimes called indirectly attributable costs. Because these costs are attributable, but not directly assignable, they are assigned to services one of two ways. In some studies, these costs are allocated across the relevant services using an allocator or cost driver that appears to be reasonably related to cost.

---

7 There are two basic tests for determining whether a cost is appropriate for regulatory purposes. The first test is called the *prudence test*, which determines whether the amount the utility spent is reasonable based on cost-minimizing criteria. There are two views. One view is that the amount spent is prudent if it was prudent at the time the decision was made. This requires accurately assessing what information management had available and used to make its decision. The other view is that the amount spent is prudent if management acted to minimize costs by fully considering changing conditions that would affect the investment. This requires assessing what management should have known and should have considered in making its decision. For example, consider a company decided to construct a power plant and, after starting construction, discovers that it would be more economical to purchase power from another company than to finish and operate the new plant. If this company finishes the new plant anyway, the costs that could have been avoided could fail a prudence test. The second test is called the *used and useful test*, which considers whether the item purchased by the utility is actually being used to provide service and that it is contributing to the provision of the service. For example, if a company has excessive numbers of telephone lines in a neighborhood, the regulator might not include some of the investment in the estimate of costs to be recovered from customers of regulated services because, even though all of the lines are used, many are not needed so they are not really useful.

8 47 CFR Ch. 1 64.901(b)(3) 10-1-05 Edition.
causation (Horngren, Sundem, and Stratton, 2005, pp. 44-46). In other instances, the costs are allocated across the activities to which they appear to be linked and that are directly assigned (Potter, et al., 2006, p. 44). Typically the allocators are measures of relative use such as the relative number of minutes or packets (Jamison, 1988).

In the second part of this step, the analyst allocates costs that appear to be common to all services across. In accounting these are called common or overhead costs. In the accounting sense, a cost is considered a common cost if its accounts or subaccounts relate to items that are used to provide multiple services. Also in accounting, a cost is considered an overhead cost if it relates to all services that the company provides such as executive costs (Horngren, Sundem, and Stratton, 2005, pp. 570 and G2). We will call these common or overhead costs in the accounting sense accounting common costs to differentiate them from how common costs are defined in economics. Analysts typically have a great deal of discretion in how they allocate these accounting common costs.9

In most writings on cost allocations, this second step creates what is called a fully distributed cost study, but this was not the original meaning of the term in regulation. Initially, the fully distributed cost method was intended to allocate all accounting costs across all services with limited regard to cost causation, which is how fully distributed cost continues to be done today (Bolter, 1978; Jamison, 1988). This is contrary to the conventional wisdom, for example in Baumol and Walton (1973) that fully distributed cost assigns all direct costs and only direct costs, and allocates only joint and common costs.

Because in practice there are a number of acceptable methods for allocating accounting costs, fully distributed cost studies can give widely varying results even though commonly accepted guidelines for cost analysts say that the allocators must be reasonable and related to principles of cost causation. Because of this ambiguity in how the accounting costs should be measured and what might be considered appropriate results, fully distributed costing methods are considered to be arbitrary and largely unusable for addressing economic or business planning issues (Baumol, Koehn, and Willig, 1987). They do, however, provide some stakeholders with a sense of fairness and transparency that economic methods do not consistently provide (Bolter, 1978; Bonbright, Danielsen, and Kamerschen, 1988, pp. 482-483; Jamison, 1998).

Classifications of accounting costs

The accounting approach classifies its costs as variable, fixed, mixed, and step. Variable costs are considered to be those that vary in direct proportion with the level of the cost driver or allocator. Certain components of digital telecommunications equipment, which exhibit no scale

9 In some instances, cost models do not differentiate between indirectly attributable costs and common costs. See for example Um, et al. (2004, p. 14).
economies, might be an example of a variable accounting cost. A fixed cost is one that is not immediately affected by the level of the cost driver, but may be affected over time. The precise amount of time that distinguishes between variable and fixed accounting costs is left vague in accounting texts. Mixed costs, also called semivariable costs, are accounts or subaccounts that have properties of both variable and fixed costs; that is to say, some portion of the accounting costs does not vary with the cost driver, but the remaining portion changes linearly with the driver. Finally, step costs are costs that are constant over a narrow range of activity, but increase in step-wise fashion with the cost driver (Horngren, Sundem, and Stratton, 2005, pp. 44-48, 89-90; Potter et al., 2006, pp. 32-33). It is unclear how the accounting approach would classify costs that vary with output in a non-step wise and non-linear fashion.

The Economic Approach to Measuring Costs

The economic approach to measuring costs is primarily concerned with identifying the costs that are caused by a particular decision such as a consumer’s decision to purchase a unit of service or a service provider’s decision to offer a particular service. As we indicated above, the approach is necessarily forward looking so that the cost information is useful for decisions at three levels: (1) whether a new company should be created or an existing firm shut down; (2) whether a firm should add a new product line to its existing portfolio or should cease offering a product line; or (3) whether the output for a particular product should be increased or decreased. We describe the economic approach by discussing costs at each level.

Total cost

In deciding whether a firm should be created or closed, the relevant cost information is called the total cost of the firm and refers to the economic value of all the inputs consumed by the company. Recall that economic value of a resource is measured in terms of the value of the best alternative use of the resource. For example, if a firm could sell an asset for as much as $1 million, but instead decides to use the asset to produce a retail product, the economic cost of the asset is the $1 million that the firm foregoes by not selling the asset, regardless of the amount the firm originally paid for the asset and independent of the amount of depreciation that has been taken on the asset. Total cost is also called stand-alone cost for the services that the company produces (Baumol, 1986, p. 120).

---

10 Technically, economies of scale exist when a proportionate increase in the level of inputs results in an increase in the productivity of the inputs (Mas-Colell, Whinston, and Green, 1995, p. 132). More commonly in regulation, economies of scale are said to exist in situations where average costs decline with increases in output (Kahn, 1988, I p. 124). Economy of scale is one of four forms of production economies. The other three are as follows: (1) Economies of scope, which exist when it is less expensive for a single company to produce two or more products than for two or more companies to produce them (Panzar and Willig, 1981); (2) Economies of vertical integration, which exist if it is less expensive for a single company to produce both an input and the final product than to produce the two separately; and (3) Economies of density, which exist when customers are sufficiently close to each other (such as in the case of urban customers) to make their marginal costs lower than the marginal cost of the average customer.

11 Economic analysis generally imposes the condition that the firm is minimizing costs given the level of output and quality that it produces (Varian, 1992, p. 26).

12 In some instances the term total cost is used to refer to costs associated with an increment of the firm’s output. This use of the term introduces ambiguity and should be avoided.
In deciding whether to create or close a firm, the decision maker should compare the present value of the future revenues of the firm, the economic values of the options created or destroyed by the decision (Trigeorgis, 1996), and the opportunity costs of resources that will be consumed by operating the firm. The opportunity costs are generally measured in terms of the present value of future cash outflows, but they may also include the opportunity costs of not selling the assets of the firm. If these opportunity costs are not included in the measure of cost for the firm, then the decision maker should examine scenarios for the firm – for example, one in which the firm operates and does not sell its assets, and one in which the firm sells its assets and closes down – and choose the scenario that gives the greatest surplus of revenue plus real options minus cost.

The issue of total cost arises in two contexts in regulation. Its primary use is in determining the price level that the utility will be allowed to charge (Kahn, 1988, I pp. 26-54). Here the accounting approach prevails, but with adjustments that make the total cost measure reflective of economic costs. These adjustments occur primarily through mechanisms of asset valuation, financial forecasts, and estimates of the cost of capital. In some instances, regulators value assets using replacement costs in an effort to reflect the economic values of the assets. Financial forecasting uses a net present value approach to estimating the overall price level of the firm (Green and Pardina, 1999). If the regulator systematically understates or overstates the economic value of the firm, owners will adjust their expectations, and the regulator’s systematic deviation from economic cost will be reflected in the company’s cost of capital.

The other context in which total cost arises as a regulatory issue is in developing price constraints for a dominant firm that allow competition from efficient competitors without creating price umbrellas for inefficient competitors. In this case, the regulator is concerned that the minimum price the utility can charge in the competitive market, while remaining financially viable, is greater than its rival’s stand-alone cost only if the rival’s entry into and sustainability in the market improves the efficiency of the market, generally defined as lowering the total cost of supplying the market demand for all of the utility products. The typical approach is to establish price floors for the dominant firm in its competitive markets such that the firm cannot price below its incremental cost (Faulhaber and Baumol, 1988). Incremental cost in this case is the incremental cost for the entire product, which is the difference between the total cost of the company if it produces this project and the total cost of the company if it does not produce the product (Baumol, 1986, pp. 113-114). We discuss this concept next.

**Service incremental cost**

Several terms can be used to denote the incremental cost for an entire product line. These include service incremental cost (Jamison, 1988), product incremental cost (Baumol, 1979), or simply incremental cost (Faulhaber, 1975). To illustrate this concept, consider a situation in which the total cost for a firm producing two products, A and B, is $1 million, but the total cost of producing A on a stand-alone basis would be $750,000. The service incremental cost for the entire product B would be $1 million minus $750,000, or $250,000. The regulator would use the $250,000 to establish the price floor for product B, assuming that product B were the competitive product, by directing the company to establish prices for B that provide the dominant firm with at least $250,000 in revenue.
To see how a price floor based on service incremental cost ensures that a rival thrives only if it lowers the total cost of serving the utility markets, consider a rival \( R \) that has a stand-alone cost of $300,000. Further assume that the regulator establishes pricing rules for the dominant firm that allow \( R \), but not the dominant firm, to provide product \( B \). The total cost to the industry to provide products \( A \) and \( B \) is the sum of the stand-alone costs of the two products, namely $750,000 plus $300,000 or $1.05 million, which is greater than the cost of allowing the dominant firm to price \( R \) out of the market and serve both markets itself. Now consider a different rival, \( R \) which has a stand-alone cost of $200,000 for producing \( B \). If the regulator imposes a price floor of $250,000 on the dominant firm, then \( R \) is able to thrive in the market, and the total cost to the industry of producing \( A \) and \( B \) is $750,000 plus $200,000, or $950,000, which is lower than the total cost were the dominant firm allowed to price \( R \) out of the market.

**Fixed costs and volume-sensitive costs.** Service incremental costs include both *fixed costs* and *volume-sensitive costs*. A fixed cost in economics is one that does not vary with the level of output. Fixed costs come about because of indivisibilities in the production process. For example, the cost of establishing a right of way on city streets for placing cable would generally be unaffected by the amount of electricity or telecommunications traffic that the utility actually put through the cable, so it would represent a fixed cost for providing service on that route. In the very short run, all costs are fixed because they cannot be changed even if production levels change.\(^{13}\) Even in the long run, some costs may be fixed. These are called long-run fixed costs or first unit costs (Sharkey, 1982, p. 37). A license to operate a utility company in a country might be an example of a long-run fixed cost. An exception would be a country in which the cost of the license is linked to the volume of revenue or service of the operator.

All costs that are not fixed are by definition volume-sensitive costs (Jamison, 1999, p. 22). However, volume-sensitive costs may be *lumpy* in the sense that production capacity must be added in lumps because of technical constraints (Park, 1989). An example might be a tower for cellular mobile phones. Increases in usage by mobile phone customers sometimes trigger the need to make the cells smaller. Sometimes this must be accomplished by adding a cellular tower, which then provides service for a large increment of traffic before requiring the addition of another tower.

Related to the concept of fixed cost is the idea of a *sunk cost*, which is the cost of a production activity that is so specialized that it cannot be easily converted to another purpose (Sharkey, 1982, p. 37). For example, a firm that obtains a license to provide service may find that it cannot reverse those costs by being reimbursed by the government or selling the license to someone else. As such, the unrecoverable portion of the license cost would be a sunk cost.\(^{14}\)

---

\(^{13}\) *Short run* refers to a planning horizon over which some inputs cannot be changed. *Long run* refers to a situation where all inputs can be changed. As such, long run does not refer to a specific length of time, but simply an absence of a constraint that keeps the business from fully adjusting its production methods (Phillips, 1993, pp. 442-447).

\(^{14}\) Sometimes the term sunk cost is used to describe embedded costs because these are monies already spent. Other times sunk cost is used to describe monies that have not been spent, but that will be even if the company goes out of business. An example would be termination clauses on contracts. Such costs are more appropriately called switching costs, which is a general term for costs incurred to change a supplier or customer (Tirole, 1997, p. 21).
In some instances, fixed costs belong to only one service, in which case they are called *product-specific* or *service-specific fixed costs* (Baumol, 1986, p. 116). In other instances, they may be shared by multiple services. We discuss this next.

**Shared costs.** In some instances, a business enjoys *economies of scope* when it provides more than one service. Economies of scope are the cost savings that exist when it is less costly for a multiproduct firm to produce products jointly than for two or more firms to divide the production among themselves (Panzar and Willig, 1981). These economies arise because of *shared costs*, which are costs incurred to produce one product that can also be used to produce another product without increasing these particular costs (Jamison, 1999, pp. 19, 21-22). For example, the cost of burying a telecommunications cable might be the same whether the cable is used only for voice telephone service or if it is used for both voice telephone service and for Internet. The Internet service might affect the technology or capacity costs of the cable, but the burying cost may be unaffected by either service, which makes the burying cost a shared cost.\(^{15}\)

Shared costs come in two forms – joint costs and common costs. *Joint costs* are the costs of producing products that can only be produced in fixed proportions (Kahn, 1988, I p. 79). The products are called joint products because they both necessarily arise from a single production process; it is physically impossible to produce one without also producing the other. At that stage of production, the products do not have separate costs. However, it is generally the case that producers refine the products further in separate processes, which would mean that in these further stages of production the products would have some separately identifiable costs in addition to their joint costs. Consider for example the production of wheat grain and straw, originally described by Marshall (1953, pp. 388-390). The cost of growing the wheat plant is a joint cost of grain and straw because growing the plant necessarily produces both. However, harvesting the crop so as to obtain both the grain crop and the straw requires more care and expense than obtaining just one or the other. As a result, certain harvesting costs such as thrashing the grain or bailing the straw are separate activities, and each has its own service incremental cost even though they share the joint cost of growing the plant. *Common costs* are shared costs for services that are not joint products (Sharkey, 1982, p. 38).

Shared costs may belong to all of the services the firm offers or only some of the services (Faulhaber, 1975; Baumol, 1986, pp. 115-117). *Incremental shared cost* is the term assigned to a shared cost that belongs to some, but not all, of the services the company provides. For example, some consumer services may have shared costs in consumer billing, but these costs are not shared with business services. *Common shared costs* are the shared costs that are shared by all of the services of the firm. The classic example is the president’s desk, but it’s not a perfect example because the desk’s cost tends to grow with the company (Jamison, 1999, p. 22).

\(^{15}\) This illustration assumes that the cable is needed for voice if Internet is not provided, or for Internet if voice is not provided. In some instances the utility might choose to use cable only if both services are offered and would choose a different technology if only one were offered. In the case where joint production triggers a change in technology, network architecture, or both, it is generally impossible to identify shared costs in terms of costs of specific operations.
Cross-subsidies. In the economic approach to measuring costs and analyzing prices, the standards of stand-alone cost\textsuperscript{16} and service incremental cost provide respectively the maximum and minimum bounds for \textit{subsidy-free prices} (i.e., prices that do not involve a \textit{cross-subsidy}), subject to two important caveats. The first caveat is that these tests apply not only to individual services, but to groups of services as well (Faulhaber, 1975). Consider the following example adapted from Jamison (1999, pp. 115-118):

Suppose a water company serves four neighborhoods. The cost of supplying water includes the cost of a well, pumping, storage tank, and transport, which is necessary if a neighborhood receives water from a well that is located somewhere other than in the town. Assume that digging a well and installing a storage tank creates $950 in fixed costs and has constant marginal costs of $0.015 per gallon. If water is provided to a neighborhood and if the well and the storage tank are not in this neighborhood, the water must be transported to the neighborhood at a cost of $90 per mile. The neighborhoods are located at the corners of a rectangle, which has a width of four miles and a length of eight miles such that neighborhoods 1 and 2 are four miles apart and are on the northwest and southwest corners of the rectangle, respectively. Neighborhoods 3 and 4 are also four miles apart and are on the northeast and southeast corners of the rectangle, respectively. Within the relevant range of prices, the citizens in each town are willing to purchase 10,000 gallons of water regardless of the price. If all neighborhoods are served by the single company, the total costs are $3,160. If any neighborhood were to go it alone, its costs would be $1,100. If two adjacent neighborhoods (i.e., 1 and 2 or 3 and 4) were to go it alone, each pair’s costs would be $1,610. Likewise, if two non-adjacent neighborhoods (i.e., 1 and 3 or 2 and 4) were to go it alone, each pair’s costs would be $1,970. Finally, if any three neighborhoods decided to band together to supply their own well, mains, water, etc., but not include the fourth neighborhood, the three neighborhoods’ stand-alone cost would be $2,480. What revenue arrangements would be subsidy free? Only arrangements that satisfied the following:

1. Revenue from all four neighborhoods equals $3,160.
2. Revenue from any single neighborhood is greater than or equal to $680.
3. Revenue from any two adjacent neighborhoods is greater than or equal to $1,550.

\textsuperscript{16} Stand-alone costs include all of the volume-sensitive costs and service-specific fixed costs for the service or services in question, the shared costs that are needed only by the service or services in question, and the shared costs that are needed by the service or services in question plus those that may be shared with other services not included in the stand-alone cost estimate. In other words, if we were estimating the stand-alone cost of services A and B, and companies also tend to offer service C, then the stand-alone cost of A and B is as follows:

\[
\text{Usage costs of A} \\
\text{Usage costs of B} \\
\text{Fixed costs of A} \\
\text{Fixed costs of B} \\
\text{Costs shared by A & B, but not C} \\
\text{Costs shared by A or B with C} \\
\text{Stand-alone cost of A & B}
\]
4. Revenue from any two non-adjacent neighborhoods is greater than or equal to $1,190.

As this example illustrates, in situations where subsets of services provided by the firm share economies of joint production within themselves that are not also shared with other services provided by the firm, then a test for subsidy-free prices necessarily must consider whether the revenue generated by this subset of services exceeds the subset’s stand-alone cost as a group. A test that considers only the service incremental cost of each service by itself is insufficient.

The second caveat is that the firm is not subject to multilateral rivalry – which is a situation where a firm’s most efficient rivals, including potential rivals, are businesses that operate in markets where this firm does not (Jamison, 1999, pp. 89-91). In the case of multilateral rivalry, prices are subsidy free as long as:

1. The utility’s revenues equal its economic costs;
2. All subsets of the utility’s products generate revenues that are no greater than those generated by the lowest prices a competitor could charge while earning zero profits and charging subsidy-free prices for its other products; and
3. All subsets of the utility’s products generate revenues that are no less than the incremental costs they create in the economy, which may be greater than the incremental costs measured at the firm level (Jamison, 1999, pp. 118-119).

To illustrate, consider the following example adapted from Jamison (1999):

Consider the four neighborhoods in the previous example, but now suppose that the cost of transporting water between neighborhoods is $100 per mile rather than $90 per mile. The costs of various arrangements for water supply are now:

1. Any neighborhood alone would cost $1,100.
2. Any two adjacent neighborhoods together would cost $1,650.
3. Any two non-adjacent neighborhoods would cost $2,050.
4. Any two neighborhoods that are diagonal from each other (i.e., 1 and 4 and 2 and 3) would cost $2,144.
5. Any combination of three neighborhoods would cost $2,600.
6. All four neighborhoods together would cost $3,339.

The most efficient arrangement is for neighborhoods 1 and 2 to have a shared water system and for neighborhoods 3 and 4 to have a shared water system, with the two sets of neighborhoods having separate systems.

Within each system, the service incremental cost of a neighborhood is $550, and the stand-alone cost of a neighborhood is $1,100. However, these cost figures do not define the range of subsidy-free prices. To be subsidy free, the pricing arrangements must ensure that each neighborhood pays at least $700 and pays no more than $950. To see why this is true, assume that neighborhood 1 refused to
pay more than $600. This would force neighborhood 2 to pay $1,050 if it continued to share a water system with neighborhood 1. However, neighborhood 2 could join the water system of neighborhoods 3 and 4 at a service incremental cost of $950, which represents the maximum that 2 would be willing to pay if it were to share a system with neighborhood 1. This option is available for each neighborhood, which gives us the range of subsidy-free prices of no more than $950 and no less than $700.

Marginal costs

Changing a firm’s output affects both the firm’s revenues and its costs. The effect on revenue is called the marginal revenue. If the firm operates in a perfectly competitive market, then the marginal revenue is simply the price at which the firm sells the marginal unit of output because the firm’s production decision does not affect market price. If the firm instead is an oligopolist or a monopolist, then an output change affects the price the firm is able to charge, and the marginal revenue is the sum of two effects. That is to say (using an increase in output to illustrate), the marginal revenue is the sum of the additional revenue the firm receives from selling more output and the decrease in revenue that the firm experiences because of the lower price it receives for its existing output. Marginal revenue is negative or zero for an oligopolist or monopolist unless customer demand for the firm’s output is price elastic, meaning that when customers experience a price change, the percentage change in the amount they purchase is greater in absolute value than the absolute value of the percentage change in price (Brown and Sibley, 1986, p. 12).

Furthermore, a change in output affects the firm’s costs: namely its volume-sensitive costs. The change in volume-sensitive costs resulting from a small change in output is called marginal cost (Kahn, 1988, I p. 65-66). A firm is said to be maximizing its profits when it chooses an output level that equates marginal revenue and marginal cost (Brown and Sibley, 1986, p. 13).

A related concept is avoided cost or avoidable cost, which is the cost that is avoided, or can be avoided, by changing the business in some way (Larson and Meitzen, 1990). Avoided cost is generally applied to a change in the service – for example, for estimating the costs avoided because of an independent power producer providing electricity into an electric utility’s power grid. Avoided cost is often used to establish differences between wholesale and retail prices and to establish costs for universal service obligations.

Because regulators often attempt to estimate marginal costs, it is important to note that marginal cost is in part determined by the regulator’s price decision. Indeed one study of how regulation uses marginal cost information defined marginal cost this way:

---

17 A perfectly competitive market is one in which there are so many suppliers that the decision of any one supplier does not influence the market price.

18 Because this firm’s output choice affects market price, the market price will decrease when this firm increases output unless other firms decrease their output by amounts at least as great as this firm’s increase in output. This would not be expected. For a review of how oligopolies behave, see Tirole (1997).

19 For example, avoided costs are used to establish discounts from retail service prices for selling wholesale services that will be resold by rivals.
The forward-looking cost of a good is the current price that makes the expected net present value of current investment equal to 0, subject to the constraint that future prices will be set so that the expected net present value of all future investments will be 0 in any period in which investment is to occur (Footnote omitted). So it is inappropriate for regulators to estimate marginal costs based on current output and prices and then adjust prices to the marginal cost estimates. However, that is what is done in practice because of the complexities of simultaneously solving for price and cost.

Marginal costs are generally used to address issues of predatory pricing, which is a pricing strategy designed by one or more conspiring firms to cause another firm or firms to exit the market (Viscusi et al., 2001, p. 170). The strategy involves two pricing stages. In the first stage, called the predatory stage, the conspirators forgo profits to price the target firm(s) out of the market. In the subsequent stage, the conspiring firms raise their prices to capture the forgone profits plus additional profits. This need to enjoy extraordinary profits in the second stage makes predatory pricing a questionable strategy because the conspirators must find a way to ensure that the profits do not attract entry. However, if the conspiring firms could exclude rivals without lowering profits, they probably would have done so in the first stage, which would have meant that predatory pricing was never needed (Larson, 1989).

The most common standard for detecting predatory pricing is the Areeda-Turner standard (Areeda and Turner, 1975). Under this standard, a firm would be considered to engage in predatory pricing if the firm were pricing below short run marginal cost. If marginal cost information were unavailable, then Areeda and Turner would substitute in its place average variable cost.

Even though the Areeda-Turner rule is typical, it is not without its critics. Some analysts believe that it is too generous to possible predators. These critics propose using average total cost as a standard. Others argue that predatory pricing is so rare and subject to so many caveats that there should be no legal rules against it (Viscusi et al., 2001, pp. 277-279).

Berg and Weisman (1992) identify certain myths about cross-subsidization and predatory pricing. One myth is that cross-subsidies are always inefficient. They explain that cross-subsidies may be important to correct for externalities such as environmental concerns or to provide service to the poor, which society decides is worthwhile. Another myth they dispel is the idea that prices must be subsidy free in every time period in order to be truly subsidy free. They explain that the relevant measures of cost and revenue in economics are the present value of cash flows and opportunity costs, so there may be periods of time during which costs and revenues for that period give the appearance of cross-subsidy. Lastly, they explain that it is a myth to believe that regulators can always find subsidy-free prices. There may be situations in which a firm is a natural monopoly, but cannot technically develop a set of subsidy-free prices.

---

Applications of the Economic Approach

Cost Studies

There are a number of complexities in estimating economic costs.

- Economic cost models can sometimes appear as black boxes, meaning that they are hard to understand. To be useful, regulators have to trust that the cost analyst has done the right thing (Jamison, 2002).
- The time horizon must be carefully defined, as must be the size of increment over which to measure the change in costs (Kahn, 1988, I pp. 70-77).
- Engineering cost studies involve making a large number of assumptions about how a firm should operate. The number of assumptions can be imponderable, and the discretion allows the possibility that desired modeling outcomes will influence assumptions (Weisman, 2000; Jamison, 2002).

Because of these and other complexities, a number of approximations of economic costs have emerged. Most of the initial applications of economic costs in telecommunications regulation focused on the degree and form of pricing discretion that the pre-1984 AT&T should be allowed in competing against telecommunications equipment providers such as Motorola and its upstart long distance competitors, primarily MCI (Jamison, 2002). The primary question for regulators was how to constrain AT&T so that its market share was based on its efficiency relative to its rivals’ efficiencies and not on an exercise of market power. This led analysts working on economic cost studies to focus on measures of marginal cost. The approximation of marginal cost was called Long-Run Incremental Cost (LRIC), which is the estimated cost of the additional investment, capital costs associated with the investment, and other costs associated with additional consumption or use of a service in the long run (Kahn and Shew, 1987).

Capacity costing became an approach for estimating LRIC in the presence of lumpy investments. Larson and Meitzen (1990, p. 155) explain the capacity cost method this way:

In essence, the capacity cost method is measuring the advancement effect caused by a given increment of demand. Whenever an increment in demand occurs, two things could happen to capacity additions. First, planned additions to capacity may be brought on line earlier to cover the increased demand for services. The advance in the timing of the future additions causes the present value of their costs to rise and, therefore, are a part of incremental costs. Second, the increment in demand may cause even greater additions to future capacity than had previously been anticipated. Again, this causes the present value of costs to increase. Together, these two effects are commonly referred to as the advancement effect. In sum, LRIC is really the advancement effect, and capacity cost is a method of measuring this advancement effect.

---

21 There was a considerable debate during this time on the use of economic costs versus fully distributed costs. AT&T favored the use of economic cost studies. Regulators generally favored fully distributed cost Bolter (1978).
Given the cost of a facility, $I$, and the total units of usable physical capacity over the lifetime of the facility, $C$, cost per unit of the facility is $\frac{I}{C}$. If a new service requires $N$ units of this capacity, an estimate of the service’s LRIC is given by its capacity cost, $CC$:

$$CC = N \ast \frac{I}{C}$$

For example, if the total lifetime capacity of a particular facility is 50,000 units and the cost of the facility is $100,000, the cost per unit of capacity would be $2. If a particular increment in demand would require the use of 10,000 units, the capacity costing method would estimate LRIC to be $20,000 for serving the increment in demand (Emphasis omitted).

As competition evolved in telecommunications, the policy issue facing regulators changed from the issue of how much the incumbent operators should sell relative to its rivals to an issue of whether the incumbent operator should even offer the service in question. This changed the interest from estimating marginal costs to estimating service incremental costs. Total Service Long-Run Incremental Cost (TSLRIC) was the term applied to studies of service incremental cost (Larson and Parsons, 1995). In some instances TSLRIC has been called Long-Run Average Incremental Cost, Long-Run Service Incremental Cost, and Long-Run Incremental Cost–Total Service. Jamison (2002) explained the difference between TSLRIC and LRIC this way:

In theory, the differences between TSLRIC and long-run incremental cost, the cost measure that AT&T advocated in the 1960s, are that: (1) TSLRIC includes the investment and expense associated with producing the entire quantity of a service, whereas long-run incremental cost covers only a change in quantity; and (2) TSLRIC includes fixed costs caused by a service. These fixed costs – also called volume-insensitive costs – are caused by providing the service and remain constant regardless of the quantity of output produced. In practice, the difference between TSLRIC and long-run incremental cost is that TSLRIC includes fixed costs. As a result, TSLRIC can miss inframarginal costs and generally understates the costs a service actually causes.

Inframarginal costs are the sum of all of the volume-sensitive costs caused by a service. Because LRIC was based on capacity cost estimates, it would understate volume-sensitive costs if there were sufficient economies of scale to outweigh any diseconomies of scale in the volume-sensitive inputs, or overstate volume-sensitive costs if the reverse were true.

After the passage of the Telecommunications Act of 1996 in the United States, the Federal Communications Commission developed its own measure of economic costs – Total Element Long-Run Incremental Cost (TELRIC) – for pricing unbundled network elements and interconnection (FCC Order, August 1996, para. 672-702). TELRIC estimates costs for an entire, hypothetical network using the best available, forward-looking technology. It is a scorched-node approach, in that it leaves fixed the switch locations of the incumbent local exchange companies, but treats all other aspects of the network – cable locations, technology choices, etc. – as
changeable. The TELRIC approach uses forward-looking cost of capital and depreciation, and includes a markup to cover some portion of common costs.

Treatment of shared costs

There are three ways that shared costs are treated in the application of the economic approach. One method is to apply what is called Ramsey Pricing or the inverse elasticity rule, which is an approach that increases prices above marginal costs in inverse proportion to the price elasticities of demand for the services, subject to the restriction that the utility receive no economic profit (Viscusi et. al, 2001, pp. 350-353). Another approach is the Shapley Value (Shapley, 1950), which allocates shared costs in a way that ensures subsidy-free prices (if subsidy-free prices exist) and that all services share in the benefits of joint production. A deficiency of the Shapley Value is that it is difficult to calculate. The most popular method is to simply assign shared costs proportionately across all services, generally meaning that each service essentially receives a uniform mark-up above its incremental costs (Um, 2004, p. 20).

Conclusion

The debate over estimating utility costs has taken some interesting turns. From the beginning of estimating economic costs in the 1960s until the mid 1990s, the interest of the incumbent telecommunications companies was to underestimate economic costs and to keep the mechanisms for estimating costs private. The incentive to underestimate costs came from the way that regulators were using costs, namely to establish price floors for competitive services. Overly low price floors were in the best interest of incumbent operators if they were able to keep rivals at bay and make up their competitive losses with revenue from monopoly services. There was also an option value in low price floors, namely that the flexibility was always there just in case economic conditions were such that it might be in the best interest of the operator to lower prices in the competitive market below marginal cost even if the company was not allowed to make up for the lost profit in the monopoly market. Keeping cost models private served two purposes. First, it also served to preserve an option value for lowering prices in competitive markets if keeping the cost model private helped keep regulators from accurately identifying long-run marginal cost. Second, the privacy helped protect trade secrets from rivals.

Once regulators began using cost models to establish maximum prices for non-competitive services, the incentives became reversed for a period of time (Jamison, 2002). Incumbents saw it in their interests to have higher prices for the interconnection and unbundled network element services that they were selling to rivals, and so sought higher estimates of economic costs. Indeed some incumbents even advocated fully distributed cost, which was something they had strongly opposed prior to the introduction of interconnection services.

It now appears that the incentives for incumbents are becoming mixed, depending on whether incumbents are net payers or net receivers of interconnection services. This makes it hard for regulators to know ex ante whether to expect cost estimates from incumbents to be low or high compared to actual costs.
Regardless of the direction of incentives for incumbents and entrants regarding whether to estimate costs too high or too low, the onus is on the regulator to require cost estimates that are based on sound economic and engineering principles and that are clear and transparent in their mechanics. Perhaps some best practices to reintroduce would be methods that track results of economic studies to ensure that their \textit{ex ante} estimates of cost reconcile with the \textit{ex post} reality of changes in the accounting records and to use benchmarking to provide reality checks for models of hypothetical companies.
References


Jamison, Mark A. 1988. “Applying Part X Allocations to Intrastate Costs.” Presented at the Fourteenth Annual Missouri Rate Symposium, Kansas City, MO.


