Draft 12/1/92

Applying Cost Indices to Incentive Regulation

By

W. Erwin Diewert and Sanford V. Berg*

Professor
Department of Economics
University of British Columbia

Florida Public Utilities Professor
Public Utility Research Center
University of Florida

Abstract

An alternative to traditional regulation is incentive regulation which bases utility rewards (and penalties) on the performance of that utility relative to a set of peer utilities. After developing performance indices for a single product firm, the index number counterparts are derived. The measurement of unit cost and productivity indices is complicated by both the number of goods, by the durability of capital inputs, and by multiple accounting periods. In conjunction with performance indices for peer utilities, yardstick comparisons can be used to reward superior performance. A plan is presented which provides incentives for the target firm to minimize discounted costs over the time horizon of the incentive plan.

*The authors are grateful to A. Nakamura and D. Sappington for comments. A more technical version of this material is available as "Index Number Issues in Incentive Regulation," in *The Potential for Using Performance Indices to Provide Regulatory Incentives*, a Report to the New York Public Service Commission, December 1992. The views expressed herein do not necessarily represent those of sponsoring organizations.

Applying Cost Indices to Incentive Regulation

1 Introduction

Regulation under which a firm's prices are tied to an exogenous benchmark breaks the link between a firm's own costs and the prices it is permitted to charge. Consequently, incentives are restored to the firm to reduce its costs. Such reductions can translate into higher earnings for the firm, which is the source of the beneficial incentives. The cost savings induced by new incentive mechanisms make the plans self-financing. Increased operating efficiencies are shared between stockholders and ratepayers.

1.1 Approaches to Incentive Regulation

Yardstick regulation represents an alternative to current approaches to regulation. The latter, as reflected in traditional cost of service incentive regulation (COSI), is not a "cost-plus" mechanism, but it can be described as a system of command and control. While yardstick regulation is not necessarily a price cap approach, it can be grafted onto price caps or be used in conjunction with current regulatory procedures. COSI is far from being abandoned in most state jurisdictions.

In his recent appraisal of regulation, Shepherd (1992) concludes "The 1980s search for a mechanical, automatic method of 'incentive regulation' was largely illusory. In complex situations, there is no easy substitute for sophisticated, effective regulation." (p. 71) This conclusion is partly based on evaluations of straw-men proposals (especially price cap incentive mechanisms). He argues that proper use of regulatory lag can restore incentives to minimize costs. However, other policy options which he did not explore (such as the use of relative performance indices) may provide superior incentives for short and long run cost minimization.

In contrast to Shepherd, Strasser and Kohler (1989) describe COSI as a "...tool [which] is at best blunt and crude, preventing the worst abuses, but not sharp enough to encourage anything better. An incentive approach promises more." (p. 137) Later they state, "Controls can keep managers from doing specific things, but they cannot command managers to use management processes energetically and creatively to tackle the problem of more efficient operation, although improved processes are essential to improved performance." (p. 169) Comparative cost is the core performance indicator in the proposal by Strasser and Kohler. However, they recommend focusing the incentives on management by linking top executive salaries to the performance reward. The firm would be given discretion as to how to allocate the pool of incentive dollars so as to achieve the greatest short and long term cost savings. We have considered incentives in the form of bonuses (resulting in higher realized rates of return). Rewards would be obtained when the cost performance of the target form was strong relative to that of comparable companies (determined by an external cost index).

A critical element of a successful regulatory plan in which a firm's prices are tied to an external cost measure (such as the average cost of a group of "similar" firms) is the appropriate choice of the external measure.[1] Ideally, this measure will reflect an achievable target for an efficient organization. If the measure systematically overstates reasonable costs for the firm, windfall gains will accrue to the firm. If the measure systematically understates reasonable costs for the firm, the firm

[1]One limitation of this option is that care must be given to defining utility output. For example, a simple, unadjusted performance indicator of unit cost equals total cost divided by total kilowatt hours. The target utility's incentives to engage in conservation or demand-side management (DSM) programs are reduced if unit production cost is the sole performance indicator. The elimination of DSM programs decreases costs (the numerator) and increases kWh (the denominator)--leading to better performance as measured by this indicator. The inclusion in the denominator of kWh saved by DSM programs is one way to remove this bias against what could otherwise be cost-effective conservation programs. The cost of DSM would be included in the numerator in this index of the cost of electricity services.

may face long-run problems in attracting capital for investment. If the measure is a good estimate of reasonable costs on average but is highly variable, substantial risk may be introduced into the regulated industry, possibly raising the firm's cost of capital. Thus, policy-makers will compare the indicators of the target firm's performance and that of the peer group when evaluating relative performance.

1.2 Overview of the Study

This study discusses some of the index number problems that arise when performance or bonus functions are constructed for a regulated utility. In order to exhibit those equivalent forms of the bonus function in an easily understood fashion, we lay out the algebra of a one output, one input regulated utility in Section 2 below. In Section 3, we exhibit the index number counterparts to the equivalent bonus functions that were obtained in the one input, one output model. We also identify the functional forms for price and quantity indexes that are widely used in empirical applications.

Section 4 addresses the practical index number issues raised by the number of goods and the durability of capital. Nine different indicoators utilizing firm and industry data are presented. Section 5 examines the implications of multiple accounting periods for the design of reward mechanisms. If we want our multiperiod performance index to provide an incentive for the target firm to minimize discounted costs over the time horizon of the incentive plan, then it turns out that the range of admissible plans is severely restricted. Section 6 shows how the range of admissible plans can be even further restricted if we appeal to fairness considerations. Section 7 describes how relative perfromance can be caluculated.

The concluding observations of Section 8 briefly address the difficult problems involved in changing output prices so that efficiency or welfare can be maximized. Features of a possible "best" incentive plan are presented.

2 Incentive Indexes for a One Output, One Input Regulated Firm

The equivalence of several performance indexes or indicators can be explained most easily when index number problems are absent. Thus, we consider in this section the case where the regulated utility produces only one output and utilizes only one input during each accounting period t (assumed to be a year).

Let p^t denote the period t price of one unit of output sold by the utility (this price is generally set by the regulator) and let y^t denote the number of units of output sold during period t. Let w^t denote the average price paid for one unit of the input used during period t and let x^t denote the number of units of input used during period t. Thus, period t cost is c^t, defined as

$$c^t \equiv w^t x^t \qquad (1)$$

The basic ex post accounting identity for the firm in period t is

$$p^t y^t = m^t w^t x^t \qquad (2)$$

where $m^t \equiv p^t y^t / w^t x^t$ is the ratio of period t revenue to cost. Put another way, m^t is (one plus) the excess profit margin for the firm in period t so that if $m^t > 1$, the firm makes pure profits in period t; if $m^t < 1$, then the firm makes a loss in period t.

The firm is instructed to supply whatever amount of output the market demands in period t at the regulated price, p^t. The regulator usually attempts to choose p^t so that m^t ends up equalling unity. However, this is difficult to do in practice, so m^t will generally not equal one. If the firm gets to keep all (or some fraction) of the excess profits, the firm will want to maximize m^t. With p^t fixed and $y^t = D(p^t)$ determined by the demand function facing the utility, then y^t will also be fixed. Hence, using (2), maximizing excess profits, m^t, will be equivalent to minimizing cost $w^t x^t = c^t$.

We assume that the regulator has a set of penalties at its disposal which is sufficient to ensure that the target utility does not attempt to degrade the quality of its outputs as a result of the incentive plan. From the viewpoint of the users or consumers of the utility's output, their welfare or profits will be maximized if p^t is set as low as possible. However, if some fraction of the excess profits of the utility are rebated back to the consumers of the utility's output, then consumers will also have an incentive to have m^t maximized or $1/m^t$ minimized. Thus, under these conditions, consumers will want to minimize p^t/m^t. Using (2), we can rewrite p^t/m^t as follows:

$$p^t/m^t = w^t x^t / y^t \tag{3}$$

$$= c^t / y^t \tag{4}$$

where (4) follows from (3) using (1). Thus, minimizing p^t/m^t is equivalent to minimizing c^t/y^t, which is the period t average cost per unit of output or unit cost.

We can also rewrite (3) as follows:

$$p^t/m^t = w^t/[y^t/x^t]. \tag{5}$$

Note that y^t/x^t (period t output divided by period t input) can be defined as the period t productivity for the utility. Thus, to minimize p^t/m^t, we can equivalently minimize w^t (the firm's input price), and maximize y^t/x^t (the firm's productivity).

Thus, for the case of a one output, one input regulated utility, we have obtained three perfectly equivalent performance indicators: (i) p^t/m^t (the output price divided by one plus the excess profit margin); (ii) c^t/y^t (unit cost); and (iii) $w^t/[y^t/x^t]$ (the input price divided by productivity).

Unfortunately, the above performance indicators are not invariant to changes in the units of measurement for the output and the input and so they are of limited use. However, this difficulty is easily remedied: we need only compare our period t performance indicators with the corresponding performance indicators pertaining to a fixed base period, say period 0 data; these satisfy the following accounting counterpart to (2):

$$p^0y^0 = m^0w^0x^0 \tag{6}$$

Using (6), counterparts to (4) and (5) can be developed where 0 replaces t. Then, upon taking ratios of p^t/m^t to p^0/m^0, we find that

$$[p^t/p^0]/[m^t/m^0] = [c^t/c^0]/[y^t/y^0] \tag{7}$$

$$= [w^t/w^0]/\{[y^t/y^0]/[x^t/x^0]\} \tag{8}$$

The various terms in square brackets in (7) and (8) can be explained in words as follows: $[p^t/p^0]$ is (one plus) the output price growth rate going from period 0 to period t; $[m^t/m^0]$ is (one plus) the excess margin growth rate; $[c^t/c^0]$ is (one plus) the rate of growth of total cost; $[y^t/y^0]$ is (one plus) the output growth rate; $[w^t/w^0]$ is (one plus) the input price growth rate and $[x^t/x^0]$ is (one plus) the input quantity growth rate going from period 0 to period t.

Equations (7) and (8) tell us that we can equivalently minimize: (i) the output price growth rate divided by the excess margin growth rate; (ii) the growth in unit cost and (iii) the input price growth rate divided by the productivity growth rate.

In order to provide the firm with an incentive to minimize costs in period t, it is important that the period 0 reference data be fixed and beyond the control of the regulated firm in period t. Thus, the period 0 data could be firm data that pertain to a period prior to the start of the incentive plan or they could be contemporaneous industry data (excluding the data of the single regulated firm for which we are developing an incentive index).

The three functions which occur on the right and left hand sides of equations (7) and (8) could each be used as an equivalent incentive index for the regulated firm with one input, one output. In the following section, we shall develop index number counterparts to the expressions which occur in (7) and (8).

3 Index Number Counterparts to the One Output, One Input Model

For simplicity, let the regulated utility produce two outputs using two inputs during each accounting period. Figure 1 depicts the points associated with different input mixes: points A and B correspond to y^0 and y^1 output levels respectively. The input mix for period 0 is x_1^0, x_2^0. Total production cost is c^0. This is a highly stylized characterization of the production process. If x_1 is the variable input and x_2 is capital (the input fixed in the short run), there are likely to be measurement problems for determining the real world correlates for our simplified model.

3.1 Cost Indices

The calculation of total costs in period 0 is straightforward. Each point along the c^0 (isocost) equation involves the same total cost (along L_1). In technical terms, production efficiency requires that inputs be hired up to the point where the ratio of their marginal productivities equals the ratio of the input prices. Of course, it is an empirical question as to whether a particular firm is actually minimizing costs for producing the observed output level. Our performance indicators only let us compare the target firm's actual performance to the comparable groups actual performance. Presumably, a well-designed incentive mechanism results in the target firm coming closer to achieving

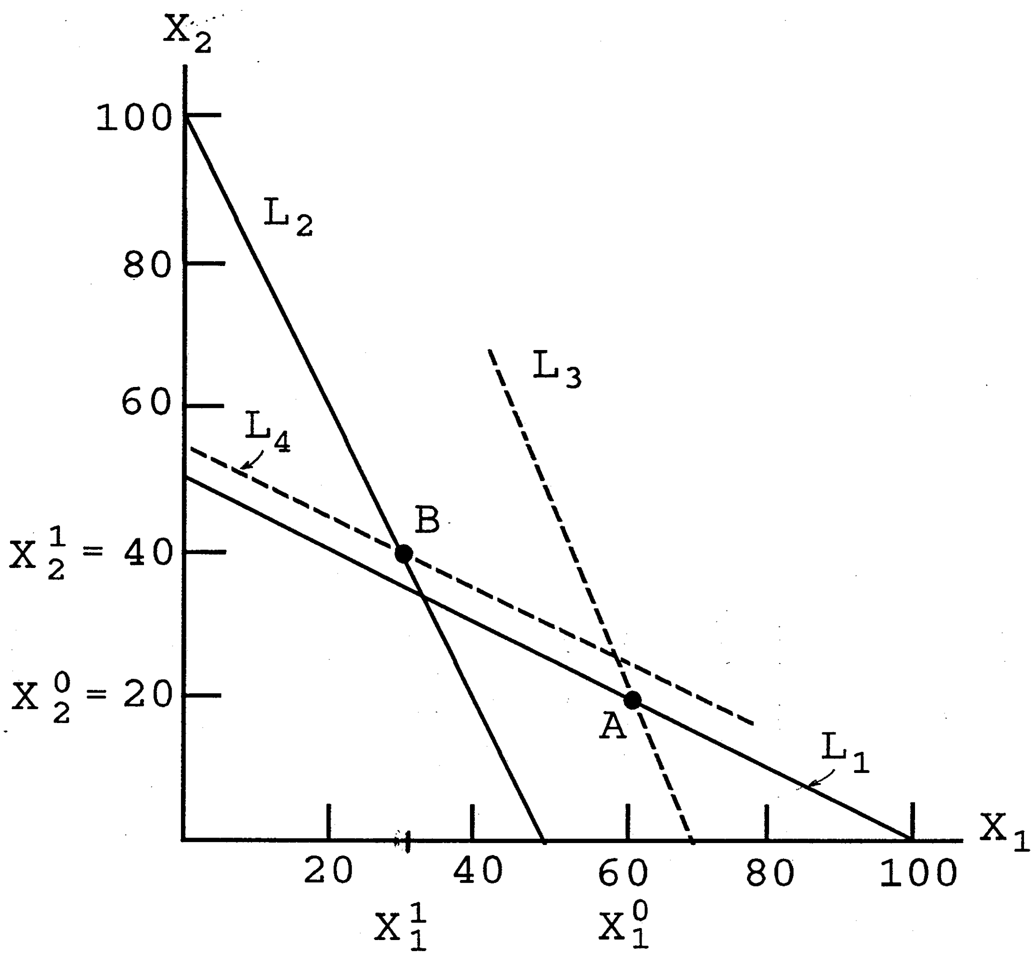

Figure 1
Index Numbers for Inputs

its potential for productivity advance--but observations on input prices and output levels do not prove that high levels of performance are being achieved.

If output in period 1 is y^1, then we can calculate the costs of producing at the cost-minimizing point again--if we know the input prices (w_1^1 and w_2^1) and the new input levels (x_1^1 and x_2^1). Thus,

$$c^1 = w_1^1 x_1^1 + w_2^1 x_2^1. \tag{9}$$

For simplicity, we assume decision-makers correctly forecast consumption levels and input prices, as well as any productivity advances. Productivity advance depends on the output levels at A and B.

Thus, if input prices changed between the periods, the calculation of the productivity advance and the cost reduction is complicated by the fact that the cost-minimizing firm will substitute away from the relatively more expensive input. The calculation of the percentage change in input prices depends on whether base period or period 1 input mixes are used.

The economic literature on this issue is vast. If we define an economic input price index as $P^*(w_1^0, w_2^0, w_1^1, w_2^1)$, then P^* can be bracketed by two price indices

$$\text{Paasche input price index} \leq P^*(w_1^0, w_2^0, w_1^1, w_2^1) \leq \text{Laspeyres input price index} \tag{10}$$

$$\text{Paasche input price index} = (w_1^1 x_1^1 + w_2^1 x_2^1)/(w_1^0 x_1^1 + w_2^0 x_2^{1y}) \tag{11}$$

$$\text{Laspeyres input price index} = (w_1^1 x_1^0 + w_2^1 x_2^0)/(w_1^0 x_1^0 + w_2^0 x_2^0) \tag{12}$$

The Paasche index uses period 1 (end period) input weights to calculate the input price index t, while the Laspeyres index uses the base period input weights. The Fisher Ideal Index is bracketed by these two.

The following example illustrates the index number problem:

$$w_1^0 = 10;\ w_2^0 = 20;\ x_1^0 = 60;\ x_2^0 = 30; \tag{13}$$

$$w_1^1 = 20;\ w_2^1 = 10;\ x_1^1 = 20;\ x_2^1 = 40; \tag{14}$$

In this case, the price of x_1 goes up, and the price of x_2 falls between the two periods. The index of the overall input price change depends on whether the period 0 or period 1 input mix is used to weight the individual prices. Consider Figure 1, where A is the period 0 mix and B the period 1 mix. The cost of producing at B is less than the cost of producing at A if the period 1 input prices are used. L_3 corresponds to the isocost line through the initial input mix but with period 1 prices: L_3 is beyond L_2.

In this example, total cost is unchanged:

$$c^0 = 10 \cdot 60 + 20 \cdot 20 = 1{,}000 \tag{15}$$

$$c^1 = 20 \cdot 30 + 10 \cdot 40 = 1{,}000 \tag{16}$$

Here, the Paasche input price index (P_p) would be

$$(20 \cdot 20 + 10 \cdot 40)/(10 \cdot 20 + 20 \cdot 40) = .8 \tag{17}$$

The Laspeyres input price index (P_L) would be

$$(20 \cdot 60 + 10 \cdot 30)/(10 \cdot 60 + 20 \cdot 30) = 1.25 \tag{18}$$

Because of the input substitution induced by the change in relative input prices, the indicator of the input price change will depend on whether P_L or P_p is used.

If, instead of inputs x_1 and x_2, we were analyzing outputs y_1 and y_2 in the base period (y_1^0, y_2^0) and period t (y_1^t, y_2^t), the very same index number issues arise. If we need a single quantity index, how is it to be determined when output prices differ between the two time periods? For example, y_1 could be residential kWh and y_2 could be industrial kWh. If the outputs have different costs of service and therefore different prices, which prices should be used to weight outputs in periods 0 and t? Functional forms can be developed for the quantity index Q:

$$Q(p^0, p^t, y^0, y^t) = p^t y^t / p^0 y^0 \; P(p^0, p^t, y^0, y^t). \tag{19}$$

For our two output example,

$$Q(p^0, p^t, y^0, y^t) = \{(p_1^t y_1^t + p_2^t y_2^t)/(p_1^0 y_1^0 + p_2^0 y_2^0)\} P(p^0, p^t, y^0, y^t). \tag{20}$$

The output price index, P, consistent with the quantity index, can take forms identified by Paasche, Laspeyres, Fisher, and others.

We turn now to the problem of choosing specific functional forms for the output price index P and the input price index P* (recall that the corresponding quantity indexes Q and Q* are determined once P and P* are chosen). There are two generally accepted approaches to this functional form determination problem: (i) the test or axiomatic approach and (ii) the economic approach. The test approach is discussed in section 3.2 and the economic approach in section 3.3 below.

3.2 The Axiomatic Approach to Choosing the Index Number Formulae

Recall that once the functional form for the output price index P is determined, then the functional form for the output quantity index Q is automatically determined. The test or axiomatic approach to the determination of the functional form for P works as follows: researchers suggest various mathematical properties that P should satisfy based on a priori reasoning (these properties are called "tests" or "axioms") and then mathematical arguments are applied to determine (i) whether the a priori tests are mutually consistent and (ii) whether the a priori tests uniquely determine the functional form for P. The a priori tests are generally mathematical analogues or generalizations of the price index in the one output case, which is the single price ratio p_1^t/p_1^0.

It turns out that the Paasche and Laspeyres indexes are subject to substitution bias. Thus, from the viewpoint of the axiomatic approach to index number theory, the choice of the Fisher ideal price index seems best. We turn now to a brief discussion of the economic approach to the problem of determining an appropriate functional form for the price index.

3.3 The Economic Approach to Choosing the Index Number Formula

Practical economic approaches to the index number problem rely on the assumption of optimizing behavior on the part of the firm. For example, if the firm is minimizing cost in each period, then the period t observed cost $c^t = w^t \cdot x^t$ will be equal to the value of the firm's period t cost function $C^t(y^t, w^t)$, evaluated at the period t output vector produced by the firm y^t and at the period t input prices w^t faced by the firm; i.e., we will have

$$w^t \cdot x^t = C^t(y^t, w^t). \tag{21}$$

The Paasche and Laspeyres input price indexes are generally biased compared to various theoretical economic input price indexes. These biases are called substitution biases and are exactly analogous to the substitution biases that occur in the consumer context when Paasche or Laspeyres consumer price indexes are used to approximate a true cost of living index. Thus from the viewpoint of the economic approach to index number theory, it seems inappropriate to use either the Paasche or Laspeyres functional forms.

Strong justifications for the use of Fisher ideal indexes can be provided from the point of view of the economic approach as well as the axiomatic approach. Moreover, the economic approach suggests that the use of Paasche and Laspeyres indexes will generally lead to biased estimates of the underlying price change. It turns out that the Törnqvist output price index, can also be given a strong justification from the viewpoint of the economic approach to index number theory.

However, if forced to choose between the Fisher and Törnqvist formulae, it would seem preferable to use the Fisher index since it has a better axiomatic justification. The economic approach to index number theory relies on the assumption of optimizing behavior and this assumption is probably not justified in the regulated utility context, since regulated firms are not allowed to choose outputs to maximize revenues; instead they are forced to supply whatever outputs are demanded at the regulated prices. Next, we turn to a discussion of three practical measurement problems that arise when we attempt to calculate index numbers numerically.

4 Practical Difficulties in the Measurement of Price Change

4.1 The Large Number of Goods Problem

It is difficult to evaluate the output price and quantity indexes if M, the number of outputs, is very large and it is difficult to evaluate the input price and quantity indexes if N, the number of inputs, is very large. In the regulated utility context, it will generally be possible to calculate output price and quantity data because the firm must collect data on the quantity sold of each output along with the associated selling price. Thus in principle, the detailed data on output prices and quantities are available somewhere in the utility's billing and accounting system.

However, in the case of input price and quantity data, the situation is quite different. Somewhere in the firm's accounting system, the purchase of each input will be recorded, but there is no necessity for the firm to decompose this individual purchase cost into separate price and quantity components. Thus when N is large, it may be very difficult (or impossible) to decompose the utility's costs into their individual microeconomic price and quantity components.

There are a number of ways of dealing with the above difficulty: (i) some quantity information may be available from various data bases other than purchases; (ii) prices could be sampled for various input categories and (iii) external price indexes could be used to deflate various categories of input cost.

4.2 Difficulties Due to the Durability of Capital Inputs

The durability of capital inputs raises another issue. The problem with calculating a price for a capital input is that the capital input lasts longer than one accounting period and hence the original purchase price of the asset should not be attributed entirely to the period of purchase. However, it is not clear how the original purchase price of the asset should be spread over the useful lifetime of the asset. Gross capital less accumulated depreciation yields net capital. If the gross and net capital calculations were empirically very close, then distinguishing which model is correct would not matter. However, it appears that in at least some regulated industries (such as telecommunications), gross capital stocks grew considerably faster than net capital stocks over the last 20 years. Thus productivity rates of growth computed using the net capital concept will be larger than those obtained using the gross capital concept. Thus, the use of historical rather than current cost accounting principles will tend to lead to lower period t costs for a target firm if the target firm has an older capital stock compared to its peer group (under conditions of inflation).

Consider also the user cost of capital. In a particular year, it consists of three components:

(1) Replacement value of plant and equipment multiplied by a real interest rate;

(2) Economic depreciation of plant and equipment;

(3) Income and property taxes.

Even though accounting versions of (2) and (3) are reported in published data, they raise conceptual issues as well. However, replacement value of plant and equipment is probably the most problematic of the three components. The proposed estimate uses the gross stock of utility investment. While this measure of capital valuation does not revalue previous investments for inflation (higher replacement costs), the omission of accumulated depreciation yields an increased estimate. Thus, the biases are more in opposing directions.

Two alternative approaches were considered for the estimation of replacement value of plant and equipment. First, calculation based on the economic theory of capital is data intensive--requiring time and effort for acquiring historical information across the set of peer firms (and the target firm). Second, the analyst can use a regulatory model of costs. Published data on the net stock of plant and equipment (based on accounting practices) can serve as the proxy for an economic calculation of replacement value. However, this approach does not provide an appropriate intertemporal allocation of capital cost since it depends so heavily on accounting conventions. The sensitivity of overall cost comparisons to the use of the gross capital stock warrants some attention. Our conclusion here is that the difficulties involved in measuring capital input prices and quantities are considerable.

4.3 The New Goods Problem

Another problem is the treatment of new goods. This problem can occur when calculating output price and quantity indexes as well as when calculating input price and quantity indexes. Consider a situation in which a new output is supplied in period t that was not supplied in period 0 or when a new input is utilized in period t that was not utilized in period 0. In both of these cases, the quantity of the new good for period 0 can be taken to be zero. However, from the viewpoint of the economic approach to index number theory, the period 0 prices for the new goods should not be zero: in the case of an input, the "correct" price for the new good in period 0 should be that price which would just induce the firm to demand 0 units of the input in period 0 while in the case of an output, the "correct" price for the new good in period 0 is that shadow price which would just induce the firm to supply 0 units of that good. Obviously, in the context of a regulated industry where the regulator sets output prices, this economic solution to the new goods problem may not make sense. Moreover, finding the correct shadow prices would involve a detailed econometric study for each new good which is too costly. Somewhat arbitrary methods will have to be used in order to deal with this problem.

4.4 Incentive Indexes that Utilize Firm and Industry Data

Figure 2 lists nine different indicators of performance and presents formulae for their calculation. The indices are based on a number of alternative definitions. The first index is based on the rate of growth of unit costs. Four indices depend on the growth rate of productivity (with and without industry input price data). Also presented are four indices relating target firm performance to peer group performance. The main advantages and disadvantages of each index are described in

Figure 2. Single Period Incentive Indices

Performance Objective	Formula*	Characterization
Own Unit Costs		
(1) Minimize Rate of Growth of Unit Cost	$I^1 = [c^t/c^0]/Q(p^0, p^t, y^0, y^t)$	Own Cost Index
Own Productivity Growth		
(2) Maximize Rate of Growth of Productivity	$I^2 = Q(p^0, p^t, y^0, y^t)/Q^*(w^0, w^t, x^0, x^t)$	Own Productivity Index
(3) Maximize Productivity Growth (Relative to Comparable Firms)	$I^3 = \{Q(p^0, p^t, y^0, y^t)/Q^*(w^0, w^t, x^0, x^t)\} - \beta t$	Own Productivity Index with Productivity Offset
(4) Maximize Productivity Growth (Industry Input Price Index Used)	$I^4 = \{P^*(W^0, W^t, X^0, X^t)Q(p^0, p^t, y^0, y^t)/[c^t/c^0]\} - \beta^t$	I^2 with Industry Input Prices and Productivity Offset
(5) Maximize Productivity Growth (Industry Input Prices, Target Input Quantities)	$I^5 = \{P^*(W^0, W^t, x^0, x^t)Q(p^0, p^t, y^0, y^t)/[c^t/c^0]\} - \beta^t$	I^4 with Industry Input Prices and Target Firm Quantities
Peer Group Indices for Yardstick Competition		
(6) Minimize Unit Costs (Relative to Comparable Firms)	$I^6 = \{[c^t/c^0]/Q(p^0, p^t, y^0, y^t)\} - \{[C^t/C^0]/Q(P^0, P^t, Y^0, Y^t)\}$	Own Costs Relative to Peer Group
(7) Maximize Productivity Growth (With Industry Productivity Offset)	$I^7 = \{P^*(w^0, w^t, x^0, x^t)Q(p^0, p^t, y^0, y^t)/[c^t/c^0]\} - \{P^*(W^0, W^t, X^0, X^t)Q(P^0, P^t, Y^0, Y^t)/[C^t/C^0]\}$ $= \{Q(p^0, p^t, y^0, y^t)/Q^*(w^0, w^t, x^0, x^t)\} - \{Q(P^0, P^t, Y^0, Y^t)/Q^*(W^0, W^t, X^0, X^t)\}$	I^3 with <u>Industry</u> Productivity Growth Offset
(8) Maximize Productivity Growth (With Industry Input Prices and Industry Productivity Offset)	$I^8 = \{P^*(W^0, W^t, X^0, X^t)Q(p^0, p^t, y^0, y^t)/[c^t/c^0]\} - \{Q(P^0, P^t, Y^0, Y^t)/Q^*(W^0, W^t, X^0, X^t)\}$	I^4 with <u>Industry</u> Productivity Growth Offset
(9) Maximize Productivity Growth (Exogenous Input Prices, Firm's Quantity Weights)	$I^9 = \{P^*(W^0, W^t, x^0, x^t)Q(p^0, p^t, y^0, y^t)/[c^t/c^0]\} - \{Q(P^0, P^t, Y^0, Y^t)/Q^*(W^0, W^t, X^0, X^t)\}$	I^5 with <u>Industry</u> Productivity Growth Offset

*Small letters refer to the target firm. Capital letters refer to industry (peer group) data: P^i = output price; Y^i = output; W^i = input price; X^i = input; $C^i = W^iX^i$ (total cost); i = 0, t -- initial period and period t.

Figure 3: Performance Incentive Indices: Intertemporal Implications

Performance Objective	Main Advantage	Main Disadvantage	Evaluation
Own Unit Costs			
(1)* Minimize Rate of Growth of Unit Cost	Does not depend on input price index or input quantity index	No allowance for input price changes	Fatally flawed if price changes Rejected as unfair
Own Productivity Growth			
(2) Maximize Rate of Growth of Productivity	Simplicity: $\frac{\text{Output Quantity Index}}{\text{Input Quantity Index}}$	Not minimize c^t	Weak long run incentives
(3) Maximize Productivity Growth (Relative to Comparable Firms)	Target utility data Adjust for input prices Allows for normal productivity growth	No incentive to reduce input prices Measurement problems, multiple outputs β^t determination contentious	Weak long run incentives
(4)* Maximize Productivity Growth (Industry Input Price Index Used)	Improvement over I^3 input price search incentives	Industry data requirements Input mix issues β^t still contentious	Long run cost minimization Riskier for target firm *INTERTEMPORALLY GOOD
(5) Maximize Productivity Growth (Industry Input Prices, Target Input Quantities)	Exogenous input prices applicable to target firm	Not minimize c^t Industry data requirements β^t still contentious	Weakened long run incentives I^5 less risky for target firm than I^4
Peer Group Indices for Yardstick Competition			
(6)* Minimize Unit Costs (Relative to Comparable Firms)	Avoids measurement problems Minimizes c^t	Industry data requirements Inflation affects index Target firm bears risks	Long run cost minimization *INTERTEMPORALLY GOOD
(7) Maximize Productivity Growth (Industry Productivity Offset)	Similar to I^3 Avoids β^t selection	Industry data requirements Target firm risk exposure	Attenuated long run incentives Not minimize c^t
(8)* Maximize Productivity Growth (Industry Input Prices & Industry Productivity Offset)	Similar to I^4 Avoids β^t selection	Same as I^4, except riskier to target firm	For fairness, equivalent to I^6 intertemporally
(9) Maximize Productivity Growth (Exogenous Input Prices, Firm's Quantity Weights)	Like I^5; less risky than I^8	Same as I^5, except riskier to target firm Not minimize c^t	Attenuated long run incentives Not minimize c^t

*Single period indices that give the target firm a clear incentive to minimize period t costs, c^t.

Figure 3, which outlines the intertemporal implications of each indicator. The final evaluation emphasizes the intertemporal properties of the index. To these we turn.

5 Incentive Indexes for Multiple Accounting Periods

Up to now, we have dealt only with incentive indexes that compare the target firm's performance in a single period t to its performance in a base period 0. For incentive schemes that cover several accounting periods, some new problems arise which are related to a problem which we have not discussed: how should the numerical value of the incentive index be mapped into a specific reward or bonus for the target utility?

If the target firm is rewarded only for good performance but not for bad performance, then the following problem could arise in the multiple period context: the target firm does well in the first period; does poorly in the second period, does well in the third period, and so on. Under these conditions, the target firm could earn a substantial bonus for a very mediocre average performance. This is the <u>cycling problem</u> discussed by Sappington (1992). More generally, suppose the target utility has the ability to defer costs from one period to the next (at an interest penalty). Then it would be very desirable to have an intertemporal, multiperiod performance indicator that remained invariant to these strategic shifts in costs.[2] This invariance can be achieved if the firm's reward function is proportional to the firm's stream of discounted costs.

A second reason for wanting the target firm's reward function to be proportional to discounted costs is that under these conditions, the target firm will have an incentive to undertake

---

[2] Mark Reeder, NYPSC, suggested this property for an incentive plan.

more rational long term investment decisions. Finally, competitive firms always have an incentive to minimize discounted costs. Thus under a proportional to discounted costs incentive plan, the target firm will be induced to behave more like a competitive firm.

In previous sections, we considered the merits and demerits of various incentive indexes that applied to the performance of the firm for only a single period t. Suppose that we want to use these single period indexes in order to construct a multiple period incentive index that will give the target firm an inducement to minimize discounted costs. Then obviously, we must restrict our attention to single period indexes that give the target firm a clear incentive to minimize period t costs, c^t. There were four such indexes that we considered: I^1, I^4, I^6, and I^8. I^1 and I^6 were to be minimized and I^4 and I^8 were to be maximized. We rejected I^1 because it would be unfair to either the target firm or its customers. This leaves us with three acceptable one period incentive indexes that were both "fair" and provided an incentive for the target firm to minimize period t costs, c^t.

Now consider a generic standard form incentive index for period t:

$$I^t \equiv a^t - b^t c^t, \quad a^t > 0, b^t > 0 \tag{22}$$

where a^t and b^t are constants which do not depend on the actions of the target firm for periods t = 1, 2, ..., T. Note that we allow a^t and b^t to depend on the target firm's actions prior to period 1. We assume that the multiperiod incentive plan starts in period 1 and lasts until period T. Let the <u>period t reward</u> for the target firm, R^t, be proportional to the standard form period t incentive index I^t; i.e., we have

$$R^t \equiv RI^t = R(a^t - b^t c^t) \text{ where } R>0. \tag{23}$$

Note that the period t reward for the target firm, R^t, will increase if c^t is decreased. Note also that R^t could be negative or positive, depending on the behavior of c^t, the target firm's total cost for period t. Under this reward function, the target firm will have an incentive to minimize the discounted sum of costs over the lifetime of the incentive plan if and only if the parameters b^t do not vary with t; i.e., if and only if we have

$$b^t = b > 0 \quad \text{for } t=1, 2, ..., T. \tag{24}$$

It can be verified that it will be fruitless for the target firm to, say, defer some costs from period 1 to period 2, provided that the interest penalty that the firm has to pay to its suppliers for these costs deferrals uses the same interest rate r^1. Thus, reward schemes of this form are <u>invariant to</u> (interest rate adjusted) <u>shifts in costs</u>. In addition, incentive plans of this type <u>provide the target firm with a clear incentive to minimize discounted costs</u> over the lifetime of the plan. Let us call a multiperiod incentive plan that has these two properties an <u>intertemporally ideal incentive plan</u>.

To ensure that the resulting incentive plan is intertemporally ideal, it will be necessary for the regulator to take away any excess profits that might be earned by the target firm in period t that are additional to the allowable amount of excess profits R^t. These excess profits would be returned to consumers as rebates that would be proportional to their period t bills.[3] On the other hand, if the regulator had set output prices for the target utility too low in period t so that utility revenues were less than utility costs c^t plus the reward R^t, then output prices would have to be raised in the following period so that the revenue shortfall could be recovered.

---

[3] Alternatively, the regulator could keep these excess profits in an interest bearing account. These funds could then be used to help offset target utility deficits that might occur in subsequent periods.

Finally, we conclude this section by noting that an intertemporally ideal incentive plan based on the period t reward function could expose the target utility to a considerable amount of financial risk if the productivity parameters β^t for I^4 were chosen to be too large or if industry costs C^t turned out to be unusually low. This risk can be limited by modifying the plan. We turn now to the problem of determining the parameter R.

6 Fairness Issues and the Determination of the Reward Schedule

If the incentive plan induces the target firm to achieve a rate of productivity advance greater than that for the peer group, there will be cost savings which can be shared by ratepayers and stockholders. As it turns out, if $R = c^0$ (i.e., the reward parameter R is chosen to equal the target utility's base level of costs), then consumers of utility services will not benefit from above normal productivity improvements and the target utility will reap all of the benefits from above average productivity improvements. On the other hand, if R is chosen to be a very small number, then consumers of utility services will get virtually all of the benefits from the productivity improvements stimulated by the above normal productivity improvements.

Of course, if R is chosen to be a very small number, then the expected reward will not be large enough to induce a productivity improving change in managerial behavior. For example, suppose there was no input price inflation so that $w^0=w^t$ for the target firm and $W^0=W^t$ for the industry, which implies that $P^*(w^0,w^t,x^0,x^t) = 1$ and $P^*(W^0,W^t,X^0,X^t) = 1$. In addition, assume that without the incentive plan, expected rates of growth in productivity were nonexistent. Suppose that the incentive plan led to a 1% increase in productivity for period t. Suppose that the parameter R were chosen to be 10% of the utility's initial cost base so that $R=.1c^0$, then the firm's period t reward as a fraction of its initial cost base, R^t/c^0, would equal .1 x {(1/1)-(1/1.01)} = .001 under these conditions. If capital costs were approximately one half of the initial costs c^0, then under the above conditions, the firm would earn an excess rate of return equal to approximately .2%. This seems rather small. On the other hand, if the parameter R were chosen to be 50% of the utility's initial cost base so that $R=.5c^0$, then making the same assumptions as before, a 1% above normal productivity improvement would lead to a 1% above normal increase in the rate of return to capital. This seems to be quite an adequate incentive to undertake productivity improvements.

Finally, fairness considerations might suggest that the benefits of above normal productivity gains be shared equally between the target utility and the consumers of utility services. This would lead to choosing the parameter R to equal one half of the target utility's initial cost base; i.e., choose R to satisfy

$$R = (½)c^0. \tag{25}$$

If the regulator felt that consumers should benefit twice as much as the target utility's owners, then choose $R = (1/3)c^0$. If the regulator felt that consumers should benefit three times as much as the target utility's owners, then choose $R = ¼c^0$. However, choosing $R/c^0 < ¼$ is probably not advisable since under these conditions, the utility's incentives to make above normal productivity improvements would be severely eroded.

The above arguments concerning the choice of R have been applied to the intertemporally ideal incentive plan where $I^t = I^{t4}$. The same fairness arguments can be applied to the intertemporally ideal incentive plan where $I^t = I^{t6}$. In both plans, our recommendation is that the parameter R be chosen to satisfy the following inequalities:

$$(¼)c^0 \leq R \leq (3/4)c^0. \tag{26}$$

7 Calculating Relative Performance

Turning to relative performance, let us consider the role the external cost index plays in determining the bonus.[4] A performance measure, X, is defined as follows:

$$X_t = Z_t^P - Z_t^N \tag{27}$$

where X_t is $-I^6$. Here, Z_t^P is the index of the change in unit cost of the peer utilities, and Z_t^N is the index of the change in unit cost for Niagara Mohawk:

$$Z_t^P = [C^t/C^0]/Q(P^0, P^t, Y^0, Y^t) \tag{28}$$

$$Z_t^N = [c^t/c^0]/Q(p^0, p^t, y^0, y^t) \tag{29}$$

Recall that capital letters refer to data from the group of comparable firms, and small letters to data for the target utility. The data are denoted by periods t and 0, where the latter is the base period. The term $Q(p^0, p^t, y^0, y^t)$ is an output quantity index based on the firm's output price and quantity vectors p^0, y^0 for period 0 and p^t, y^t for period t. The functional form for Q would be the Fisher ideal quantity index:

$$Q(p^0, p^t, y^0, y^t) = [p^t \cdot y^t/p^0 \cdot y^0]P(p^0, p^t, y^0, y^t) \tag{30}$$

Here, the price index P is Fisher ideal price index:

---

[4]Mark Reeder provided useful background for this section.

$$P_F(p^0, p^t, y^0, y^t) \equiv [(p^t \cdot y^0/p^0 \cdot y^0) \cdot (p^t \cdot y^t/p^0 \cdot y^t)]^{1/2} \tag{31}$$

The Törnqvist price index could also be used. The formulas are presented so readers can see how the calculation of cost per unit requires appropriate treatment of period 0 and period t data.[5]

Returning to X_t, the difference between peer group and target firm unit cost increases, we see in Figure 4 how the reward (or penalty) depends on the level of cost savings associated with each percentage point difference in the cost change for the target firm relative to the peer group. Beyond the index number issues described earlier, we need to address the incremental incentives facing the target firm for improving its performance.

If we assume no inflation, what are the relative savings associated with a cost increase of Z_t^N instead of Z_t^P? If the target firm's unit costs rise by 3% but unit costs rise by 5% for peer utilities between t = 0 and t = 1, then the target firm outperformed the comparable group by two percentage points. Thus, the total cost saving (S_t) associated with the superior performance can be approximated by

$$S_t = .02 \cdot c^0 \tag{32}$$

where c^0 is the target firm's base period costs.

[5]Note that environmental costs are not included in the unit cost calculation. If policy-makers have a high degree of confidence in valuations placed on water, air, and surface emissions, these could be incorporated in the performance indices.

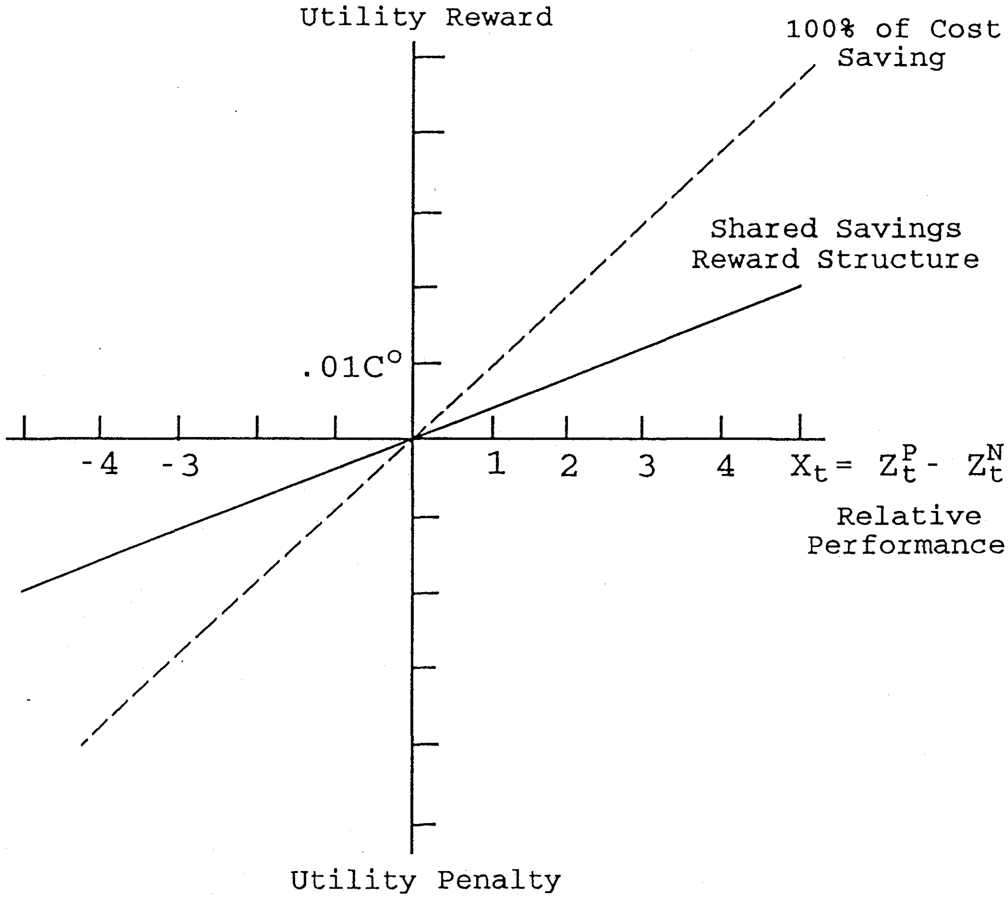

Figure 4
Reward Structure for Relative Performance

We have already noted how these savings can be translated into a reward (or bonus) for the firm. Part of the cost savings are translated into customer prices which are lower than would have been experienced under a 5% increase in costs. If half the savings were retained by the shareholders, price would rise by only 4%.

Alternatively if peer costs fell by 2% and the target firm's costs were unchanged, the firm would be penalized by a 1% lower price (under a 50-50 sharing rule) for the next period. Note that lump sum charges (or rebates) per customer would avoid introducing problems associated with elasticities and would interfere less with the role of price signals in reflecting incremental costs. Program designers would need to examine this aspect of the incentive mechanism.

We note that in both of our intertemporally ideal incentive plans, the structure of the target utility's output prices is held fixed; i.e., for every period t in the incentive plan, the target utility's output price vector p^t is proportional to the base period price vector p^0. Section 8.1 briefly considers the problems involved in changing the structure of output prices.

8 Concluding Observations

Despite the fact that classical Cost of Service Incentive (COSI) regulation has been able to adapt to dramatic economic changes (including input price instabilities, environmental regulations, and competitive pressures), the ad hoc adaptations have produced a complex layering of rules and cost allocations for electric utilities. The achievement of regulatory objectives may be thwarted by rules which are potentially inconsistent with one another. In particular, the incentives for cost reduction are attenuated by relatively short regulatory lag under COSI. Managerial efforts which have long run payoffs (in terms of production efficiency gains) are particularly problematic.

Furthermore, concern over achieving the appropriate level of energy conservation activity has added new layers of regulation. Because of these developments, we focused on an alternative incentive mechanism which takes advantage of regulatory experience and theoretical developments over the past two decades.

8.1 Efficiency and the Structure of Output Prices

Our suggested intertemporally ideal incentive plans have at least one major defect: they leave the relative prices of the target utility's outputs frozen at their reference period levels. Thus, our suggested incentive plans do not address the problem of finding "optimal" output prices in each period that would maximize economic efficiency (or welfare). It is often argued that in order to achieve an efficient allocation of resources, the regulator should force the target firm to set its output prices equal to marginal costs. However, in regulated industries, typically prices are above marginal costs. Hence if the regulator attempts to implement the marginal cost pricing rule, this will involve lowering the target utility's output prices until they equal marginal costs. There are at least three problems associated with attempting to implement this marginal cost pricing rule. The first problem is that typically neither the regulator nor the target utility will know with any degree of certainty what the appropriate marginal costs are. The second problem is that as soon as output prices are lowered, the target utility may not make enough revenue to cover costs (and its reward). Hence the extra revenue has to come from somewhere and typically some deadweight loss in the economy will be generated by the attempt to raise this extra revenue (unless multipart pricing is utilized). The third problem is that the regulator will not want to become directly involved in setting output prices since consumers who are hurt by regulator induced price changes will loudly complain about the unfairness of the price setting process.

Nevertheless, there is a need for the regulator to get involved in the price setting process. How should new target utility products be priced? Should the target utility be allowed to lower its prices to meet the prices of actual or potential competitors? These questions and others require a scientific framework for setting output prices in a regulated industry. Unfortunately, such a framework does not yet exist or has not yet found widespread acceptance. This explains why we did not address the thorny problems of the optimal pricing of outputs in our suggested incentive plans: a coherent acceptable framework for output pricing does not yet exist.

8.2 A Possible "Best" Incentive Plan

Returning to the proposed external cost index, the main conclusions can be summarized as follows:

(i) The Fisher functional form P_F for the price index should be used when constructing regulatory incentive indexes.

(ii) The incentive plan should be intertemporally ideal; i.e., the target firm's overall reward should be unaffected by (interest rate adjusted) shifts in costs and the plan should provide the target firm with a clear incentive to minimize discounted costs over the lifetime of the plan.

(iii) We have suggested two intertemporally ideal incentive plans defined by specifying a period t reward $R^t=RI^t$ where in one case, $I^t \equiv I^{t4}$ where I^{t4} reflects relative productivity growth and in the other case, $I^t \equiv I^{t6}$ where I^{t6} reflects relative unit costs. The parameter R should satisfy the bounds given by (85). R can be chosen so that the target firm and its consumers share the benefits of above normal productivity improvements equally.

(iv) In choosing between the two suggested intertemporally ideal incentive plans, we would lean towards the use of the second plan which has $I^t = I^{t6}$, since in this case, the regulator does not have to pick a sequence of expected productivity growth rates β^t as is the case with the first plan. The problem with the first plan is that if it becomes evident that an inappropriate sequence of productivity growth rates β^t were picked initially, then pressures will emerge to renegotiate the β^t.

(v) The intertemporally ideal incentive plans can be modified so that the target firm will never have to pay a penalty. This probably should be done to limit the financial risk of the target utility.

(vi) The optimal structure of output prices remains to be determined. In our suggested incentive plans, output prices are adjusted proportionally to achieve a period by period balance between target utility revenues, costs and rewards.

The potential for yardstick regulation seems clear. Incentives for cost reduction would be enhanced if the target utility faced the penalty and reward structure described here. If an appropriate comparison group can be identified, consumers can benefit from a movement away from cost of service regulation.