Second-Generation Incentive Regulation for Ontario Power Distributors

Pacific Economics Group

Economic and Litigation Consulting

Second-Generation Incentive Regulation for Ontario Power Distributors

13 June 2006

Mark Newton Lowry, Ph.D.
Partner

PACIFIC ECONOMICS GROUP
22 East Mifflin, Suite 302
Madison, Wisconsin USA 53705
608.257.1522 608.257.1540 Fax

Table of Contents

I. INTRODUCTION

The Ontario Energy Board is in the process of developing a second-generation incentive regulation plan for provincial power distributors. The plan is envisioned to be of price cap form. The maximum duration of the plan is three years. Following a rate case, companies completing the plan will embark on a third generation incentive regulation regime that is also under development.

Pacific Economics Group ("PEG") was recently retained by the Board as an advisor on incentive regulation, also referred to as performance-based ratemaking ("PBR"), issues. PEG is the leading North American consultancy on PBR for gas and electric utilities. The design of rate escalation mechanisms and other plan provisions is a company specialty. We previously provided advisory support to the Board for the Natural Gas Forum. We have also advised several Canadian utilities and trade associations.

In the next section I establish criteria for the design of regulatory systems and use this framework to discuss PBR and its advantages over the traditional cost of service approach to regulation. I then consider at some length the major issues in the design of a PBR plan. This discussion is intended to provide stakeholders with a common frame of reference in considering the issues that arise in PBR plan design. The paper concludes with recommendations for second generation incentive regulation for power distributors in Ontario.

II. PERFORMANCE BASED REGULATION

1 RATIONALE FOR PBR

PBR is a well-established alternative to traditional regulation of energy utilities. In North America, PBR has now spread from such "nursery jurisdictions" as California, Massachusetts, and British Columbia to a much broader group of jurisdictions that includes Alberta, Florida, Iowa, North Carolina, Ontario, and Oregon. The Federal Energy Regulatory Commission (FERC) and Canada's National Energy Board (NEB) use PBR to regulate oil pipelines and some gas lines. PBR is also extensively used in North America to regulate railroads and telecommunications utilities. Overseas, PBR is even more ubiquitous and is, effectively, the standard means of regulating for investor-owned utilities. Due in part to the many jurisdictions using PBR and the varied industries involved, diverse approaches have developed. This means that many established mechanisms are available today to craft a PBR plan.

In this section of the report, I propose a sensible set of criteria for the design of regulatory systems. This is used to develop the rationale for PBR and identify its comparative advantages vis-à-vis traditional regulation.

1.1 System Design Criteria

1.1.1 Efficiency

One of the most important criteria for evaluating alternative regulatory systems is their ability to promote economic efficiency. A regulatory system is (economically) efficient to the extent that it generates the maximum possible net economic benefits for society. Economic efficiency has several dimensions. One is the operating efficiency of the utility. This has marketing as well as a production dimension. The cost efficiency of the regulatory system also matters. I discuss here the concepts of production efficiency, marketing efficiency, and regulatory cost efficiency in turn.

Production Efficiency

Regulation encourages a utility's production efficiency to the extent that it induces it to produce the services that it provides at minimum cost. In the short run, capital inputs are substantially "fixed" in the sense that adjustments in the amounts used are quite expensive. Productive efficiency then depends primarily on the extent to which services are provided with a minimum-cost mix of other, variable inputs such as labor. In the long run, all inputs are variable and the cost-effective use of capital is also an efficiency concern.

Marketing Efficiency

Regulation encourages a utility's marketing efficiency to the extent that it induces it to provide the right mix of services to the right customers in the right amounts. Simply put, we want utilities to play the right role in our evolving economy. The right role is one in which they help households, business, and other customers pursue their goals at the lowest cost. A good measure of the success of marketing is the difference between the value of services to customers and cost of service provision. In the short run, the adjustment in the rates and other terms of existing services to reflect changing market conditions is the main marketing challenge. In the long run, the mix of services offered by a utility becomes an important concern.

Service quality is an important aspect of the terms of service. Customers care about the quality of services as well as their prices. Customers' need for quality varies and changes over time. Unregulated markets often involve an array of products with different price-quality attributes. Since social benefits from regulation depend on both price and quality, the encouragement of appropriate quality levels is a proper regulatory objective.

The marketing efficiency of a utility does not depend solely on the terms on which it offers services to markets that, due to the essential character of the services and the lack of competitive pressures, are regulated. Utilities may also be able to enhance welfare by supplying customers in unregulated markets. These are sometimes referred to as "non-core" markets. Almost every utility has some involvement in such markets. The rental of underutilized real estate under transmission lines is a good example. Utility participation in unregulated markets can lower prices and make valuable new products available to customers. These advantages are especially attractive in markets, like those for local telecom services, where additional competition is needed.

<u>Regulatory Cost</u>

Costs are incurred in utility regulation. These include, most obviously, the resources (*e.g.* lawyers, accountants, engineers, and facilities) of utilities, intervener groups, and government agencies that are dedicated to the regulatory process. Senior company officials are also drawn into the regulatory arena. This can impair utility performance to the extent that it distracts managers from their operating responsibilities.

1.1.2 Fairness

A second fundamental criterion for appraising regulatory systems is fairness. This concerns the manner in which the benefits of utility operations are divided among the stakeholders in the regulatory process. A minimum condition is that the chief parties to regulation, shareholders and customers, fare no worse under PBR than they would under traditional regulation. A more aggressive standard would be for the chief parties to share in the benefits of improved performance that PBR makes possible.

In assessing the fairness of the regulatory system, it is important to remember the outcomes that matter to stakeholders. Customers benefit, most obviously, from low prices and high service quality. Customers also benefit from rate stability and the availability of tailored rate and service offerings.

1.1.3 Conclusion

Regulation should encourage good utility performance, use regulatory resources efficiently, and share the benefits of good performance between utilities and their customers. Utility performance has a marketing as well as a cost containment dimension. Good performance and a fair sharing of benefits both point to the need for a proper balance between a utility’s operating risk and expected return.

1.2 The Regulatory Challenge

1.2.1 Cost of Service Regulation

Description and Precedent

Cost of service regulation ("COSR") is a convenient term for the traditional approach to the regulation of North American energy utilities. Under this system, the rates approved by a commission are expected to recover the company's prudently incurred cost of providing regulated services. This cost includes a return on capital.[1] Rate cases are held periodically in which estimates are made for a certain test year of the prudent cost of capital, labor, and other inputs that are used to provide regulated services. This becomes the base rate revenue requirement.[2] [3] The test year may be a future or a recent historical year.

Once the revenue requirement is determined, it must be allocated for recovery from tariffed services. The rate charged for each service recovers this assigned cost given data on peak demand, delivery volumes, the number of customers, and other billing determinants. The regulated service offerings and rate designs require commission approval. Reviews of these issues do not always coincide with rate cases.

Evaluation

COSR has played a vital role in the development, in several countries, of utility industries that make service widely available at an affordable cost. Its focus on the cost of service has two cardinal benefits. One is the satisfactory resolution of the issue of fairness: the utility has a fair chance of recovering its cost of service if it is well-managed but has only a limited opportunity to earn more. The other advantage of COSR is the reduction in utility operating risk that results. This ensures that capital can be obtained for utility undertakings at a reasonable price.

---

[1] This characterization of cost of service regulation is, of course, stylized. The terminology and precise procedure for setting rates under COSR varies considerably across regulated industries and regulatory jurisdictions.

[2] The volatility of energy prices has prompted some regulators to provide for a shorter lag between the purchase of energy inputs and the addition of these costs to the revenue requirement.

[3] The determination of allowed cost is complicated if the utility company sells some products in non-core markets. To the extent that a utility serves such markets, its total cost will exceed the cost of the regulated services that it provides. In that case, some share of total cost must be assigned to non-core services. Alternatively, the revenue obtained from such services may be netted off of total cost.

The recovery of capital cost is especially important in the assessment of utility risk. Capital goods provide services over many (*e.g.*, 30-40) years. Once capital goods are installed and become utility plant, their value in alternative applications is often well below their cost. Companies owning such "relationship-specific assets" are vulnerable to opportunistic changes in the compensation allowed by regulators[4]. COSR is a good basis for a regulatory compact that reduces the likelihood of such outcomes.

These benefits of COSR help to explain the sizable scale on which COSR has been used. It was, for many years, the standard means by which investor owned utilities were regulated. The chief alternative to COSR in those years was state enterprises. Large utility industries were built under COSR, including those in the United States and Japan. In Canada, state and municipal enterprises dominated the industry for many years. However, most of these companies have now been corporatized and subject to COSR for several years now. COSR has also been prevalent in the regulation of natural gas utilities.

Despite this lengthy track record of effectiveness, there is mounting evidence that COSR does not always achieve the maximum net benefit to society that is achievable from utility services. One fundamental problem is the high cost that must be incurred for regulators to learn about utility operations. If they understood the changing constellation of production and marketing practices that are ideal for the situations of specific utilities over time, they could in principle mandate the services that should be provided and their terms. Unfortunately, it is difficult even for experienced managers in an industry to recognize best practices given the uncertainty that exists regarding future supply, demand, and policy conditions. The challenge is much greater for regulators and customers who lack operating experience in the industry. Economists call this situation one of information asymmetry. A redressing of this asymmetry requires substantial exchange, processing, and analysis of information.

Another fundamental challenge in COSR is the allocation of the common costs that utilities incur in providing miscellaneous regulated services and any non-core services. The inherently arbitrary nature of common cost allocations makes them a source of controversy.

Measures are naturally taken to contain the cost of COSR. One option is to reduce the frequency between rate hearings. Another is to scale back the scope or intensity of prudence

---

[4] Utilities may also be denied recovery of prudent operating expenses but can at least suspend purchases of operating inputs once a pattern of denial becomes apparent.

reviews. For example, companies may be placed at significant risk only for actions with conspicuously unfortunate outcomes. The extent to which utilities fall short of best operating practices is rarely considered.

Regulatory cost can also be contained by restricting practices that complicate regulation. For example, companies may be discouraged from offering diverse services or complex, changing rate structures. They may also be discouraged from engaging in practices that are novel, risky, or inherently controversial.

All of these measures can reduce regulatory costs. Unfortunately, some of these economy measures can also compromise the production and marketing efficiency of utilities. To the extent that prudence reviews are limited, for instance, rate adjustments tend to reflect the trend in a utility's own unit cost. Efforts to trim costs or to improve the market responsiveness of rates and services then lead eventually to lower rates. This weakens utility performance incentives. Another class of initiatives that is strongly discouraged is those involving a significant risk of conspicuous failure. This would include many kinds of innovations.

Restrictions on utility operations that are hard to regulate can also reduce efficiency. For example, limited and inflexible rate and service offerings hamper a utility's ability to satisfy customers' complex and changing needs. Customers may not use utility services even when they can be provided at a lower cost than the available alternatives.

The efficiency consequences of ineffective marketing are especially acute where demand is elastic (sensitive) with respect to rates and other terms of service. Situations in which demand is elastic include those in which customers can obtain their service needs in other ways at a competitive cost. Demand is also frequently elastic for incremental uses of utility services by existing customers. A third important source of demand elasticity is economically distressed businesses that make extensive use of utility services. When elastic customers do not make optimal use of a utility system, the margins from services to them are lower than they can be and a larger share of the utility's common cost must be recovered from other services.

An important special case of demand elasticity is that in which the market for a service has a structure that is so competitive that regulation is unwarranted. Such markets are called non-core markets earlier in this paper and in discussions of utility policy. Utility companies sometimes seek to achieve economies of scale and scope by diversifying into non-core markets. These markets are sometimes served by the utility itself. A familiar example is the rental of land in power transmission corridors. Alternatively, the economies can sometimes be captured by an affiliate that serves non-core markets and also performs certain operating functions for the utility.

Under COSR, utility service to non-core markets and purchases of services from affiliates can both raise cross-subsidization concerns that raise regulatory cost. Some regulators have responded to this challenge by limiting non-core market offerings by utilities and affiliate transactions or by establishing onerous rules that effectively discourage them. These problems and the attendant regulatory costs may lead utility companies either to forgo non-core market involvement or to serve these markets through unregulated affiliates that do not meet customer needs as well and/or fail to exploit the full potential benefits of scale and scope economies that are available from consolidated operations. Failures of unregulated affiliates of utilities are reported routinely in the trade press, and many of the failures reflect this problem.

This outcome is unfortunate since customers in non-core markets can benefit from utility participation, as noted above. Non-core market involvement can also permit some sharing of common costs and permit the utility to offer better terms of service to core customers.

One economy measure that can increase the efficiency of COSR is a reduction in the frequency of rate cases. As the period between rate cases, sometimes called regulatory lag, lengthens the period of time during which the company retains the benefits of performance gains increases. Performance incentives are thereby strengthened, especially for projects with longer pay back periods. This also makes it easier for regulators to afford utilities greater marketing flexibility.

The ability of a utility to operate without rate adjustments depends critically on the extent that its unit cost exhibits a rising trend and tends to fluctuate around its trend. Unit cost is more likely to exhibit a flat or declining trend when input price growth is slow and utilities are able to realize rapid productivity growth. Unit cost is more likely to exhibit stability around its trend to the extent that input prices and demand have stable trends, cost is not subject to major external shocks for other reasons, and investments tend to be spread evenly over time.

The unit costs of energy utilities, unfortunately, tend to rise over time and are sometimes volatile. Unit cost volatility is especially pronounced in the procurement of energy inputs. The chief reason for rising unit costs is that productivity growth in the energy utility industries, as in most sectors of our economy, cannot keep pace with input price inflation.

The end result of these conditions under COSR is that rate case cycles in the energy utility industries typically do not exceed three years. Recovery of fuel and purchased power costs often occurs even more rapidly using special fuel adjustment clauses. This situation is not conducive to strong performance incentives and limits the operating flexibility that regulators are comfortable granting.

1.2.2 The PBR Alternative

The term PBR applies to a variety of regulatory mechanisms and procedures that differ from COSR in relying less on a utility's own cost, output, and service quality to establish its rates and other terms of service.[5] An economist might call the resultant decoupling of rates from a utility's own operating data an externalization of the regulatory system. Externalization can be achieved in several fundamentally different ways. One is to reduce the frequency of rate cases which, as we have seen, cause a company's rate trend to more closely match its unit cost trend. Another is to avoid a complete unit cost true up when such adjustments are made.

There are several "active ingredients" in this new approach to regulation. One is automatic rate adjustment mechanisms that are established in advance of their operation. Such mechanisms are often represented by mathematical formulas. The use of such mechanisms can reduce the frequency and scope of regulatory interventions that would tend to make a utility's rate trends more similar to its unit cost trend. A second source of progress is a reliance for ratemaking purposes on data that are insensitive to the actions of utility managers. Data that are useful in this regard include indices of price inflation and information on the operations of other firms in the industry.

Economic research is a third active ingredient of PBR. Theoretical and empirical research can be brought to bear on the appropriate combination of automatic mechanisms and external data. One example is research on external rate adjustment mechanisms that yield revenues sufficient to compensate a competently managed utility. Another is research on what plan provisions provide balanced and strong performance incentives.

To the extent that such tools permit rate caps and other limits on terms of services to be established by external means, the limits are insensitive to a utility's own performance improvement initiatives. Utilities will then find that their performance has a greater impact on earnings. This strengthens incentives to improve operating efficiency. The externalization of the rate setting process also lessens concerns about cost shifting and cross subsidies. With stronger incentives and lessened ability to shift costs, utilities can be given more operating flexibility.

The combined effect of these attributes is a regulatory system that, in many cases, can stimulate better utility performance despite lower regulatory cost. PBR can thus increase the size of the economic "pie" that is available for sharing between utilities and customers. It therefore constitutes an advance in the "technology" for utility regulation.

While results to date have been encouraging, the state of the art is not so far advanced that PBR is markedly superior to COSR in all cases. One problem area is risk. Compared with COSR, PBR will often expose utilities to more conventional business risk. Their situation in this regard is much like an airline that, faced with soaring jet fuel prices, can hope for some relief from market-based fares but is by no means ensured full compensation. There is, additionally, a greater regulatory risk that restrictions on rate and service offerings will be established in an arbitrary manner that denies a well-managed utility a reasonable chance of recovering its cost. The recovery of capital cost is a particular concern since the utility has less flexibility to reduce these costs if the terms of a rate plan are unreasonable.

Any increase in utility operating risk under PBR will ultimately be recognized by capital markets and reflected in the cost that utilities incur to attract funds. The increase in the cost of funds can significantly erode the net benefits of PBR. It can also cause utilities to oppose PBR if they feel that it does not offer a reasonable balance of risk and return. One consequence of this general problem is that PBR still involves occasional trueups of a utility's rates to its cost.

Our analysis suggests that the advantages of PBR over COSR depend on the particulars of its application. PBR will generally be more advantageous to the extent that effective COSR is unusually costly. For example, when the unit cost of utilities is rising due to brisk input price inflation, frequent rate cases are required to compensate utilities under COSR whereas PBR can offer automatic inflation adjustments. COSR can also be unusually costly, as we have seen, when rate cases involve unusually difficult issues of cost allocation, transfer pricing, or operating prudence. Consider, lastly, that COSR is unusually costly when regulators have jurisdiction over a large number of companies and/or face a sudden and sizeable increase in the number of

---

[5] Other names for this approach to regulation that are sometimes used include incentive regulation and alternative regulation (Altreg).

jurisdictional companies. PBR can reduce the cost of regulation in this situation to more manageable levels.

PBR will also be more advantageous to the extent that the effective mechanisms that have been developed to date are amenable to implementation. As discussed further in the next section, for instance, a common approach to rate indexing requires good estimates of historical productivity trends of utilities. The calculation of productivity indexes requires, in turn, a considerable amount of historical operating data. Estimates of the historical productivity trend must, furthermore, be reasonably good estimates of *future* productivity trends. These conditions do not hold in every possible application. For example, good historical data may be unavailable and past productivity trends may not continue in the future.

A third set of circumstances that affects the advantage of PBR is the opportunities for utility performance gains. The extent of performance gains achieved, after all, will depend in part on the performance gains that can be achieved by stronger incentives and better operating flexibility. Generally speaking, the potential for performance gains is greater to the extent that more of the activities that contribute to performance can be controlled by utility personnel. The potential for performance gains is also larger to the extent that subject utilities are substandard performers and/or can improve performance with better operating flexibility.

A quick review of the jurisdictions where PBR is prevalent around the world reveals that it does tend to be more widely used in situations where advantages such as these are larger. For example, PBR is especially prevalent in activities that are difficult to regulate under COSR and require operating flexibility. Most notably, it is the standard approach to the regulation of railroads, oil pipelines, and telecom utilities, which need substantial marketing flexibility if they are to serve markets that have varied competitive pressures from a common set of assets. In the United States, PBR is also widespread in the regulation of natural gas procurement, which involves a price-volatile input and difficult issues of operating prudence.

It is also interesting that PBR is the standard approach to the regulation of newly privatized utilities. Such privatizations are widespread outside North America. Decades of operation as public enterprises make it likely that such utilities are capable of unusual short-term performance improvement.

There is also some evidence that PBR is more common where regulators have jurisdiction over a large number of companies. For example, PBR is less common in North

America than overseas. Most energy utility regulation in North America occurs at the state and provincial level. Regulators in most state and provinces don’t have jurisdiction over a large number of utilities.

III. PBR PLAN DESIGN

2 RATE CAPS

Most approved PBR plans to date have involved multiyear caps on the growth of utility rates or revenues. This section addresses the rate cap, also referred to as the price cap approach. This approach generates stronger incentives to improve marketing performance. Greater marketing flexibility is sometimes allowed. Marketing flexibility provisions of rate cap plans are discussed at some length. The following section addresses the alternative revenue cap approach.

2.1 Overview

Under a rate-cap plan, restrictions are placed on the escalation of rates for utility services. The restrictions can be placed on annual rate escalation or on the cumulative escalation during the plan period. The limits are called caps since utilities are usually free to charge rates that are less than the maximum allowed under the escalation restrictions.

The mechanisms for limiting rate growth are diverse, but all have the attribute of being external to the company's operation. The simplest approach is to hold rates constant for the plan duration. This approach is called, variously, a rate freeze or rate case moratorium. A simple variant of the rate freeze --- sometimes called a "stair-step" mechanism --- is a sequence of pre-scheduled rate adjustments, which may be increases or decreases.

Rate growth can, alternatively, be capped using indexes. Under this approach, growth in baskets of the utility's prices may be measured using actual price indexes ("APIs"). Growth in each API is limited using a price cap index ("PCI").[6] Here is a mathematical rule for limiting the growth in annual rate escalation:[7]

$$\text{growth } API \leqq \text{growth } PCI \,. \qquad [1]$$

[6] The useful acronyms API and PCI appear to have developed in US Federal Communications Commission proceedings.

[7] A mathematical rule for the *cumulative* escalation of rate is:

$$API_t \,/\, API_o \leq PCI_t \,/\, PCI_o.$$

Here API_o and PCI_o pertain to the base period.

Price cap indexes are largely external to the company's operations. Their growth is typically determined by a formula that includes a price inflation measure. The design of such indexes is discussed further in Section 5 below.

2.2 Rate Cap Precedents

2.2.1 United States

Rate Indexing

In the United States, the first large scale use of rate indexing was a plan for Class I Line Haul Railroads under the terms of the Staggers Rail Act of 1980.[8] An index was used to adjust a zone of rate freedom in which rates to captive shippers were free from challenge. The US telecommunications industry was another rate indexing pioneer. The Federal Communications Commission (FCC) played a leadership role in this regard, approving rate cap plans for AT&T in 1989 and for interstate services of local exchange carriers (LECs) in 1991.[9] Index-based rate caps are now widely used in state-level telecom regulation.

In the US energy industry, rate indexing has been featured in PBR plans for several utilities. The first rate plan with indexing for a US electric utility was that for the bundled power services of PacifiCorp (CA). Since then, plans have been approved for the bundled power service of Central Maine Power (ME) and the power distribution services of San Diego Gas and Electric ("SDG&E") and Southern California Edison (CA), Bangor Hydro Electric and Central Maine Power (ME), National Grid and NSTAR Gas and Electric (MA), and Narragansett Electric (RI).

Massachusetts regulators first approved a rate indexing plan for Boston Gas in 1996. Since then, they have approved a new plan for this company as well as plans for Bay State Gas and Berkshire Gas. Plans with rate indexing have also been approved for gas delivery services of SDG&E (CA), Bangor Gas (ME), and Transwestern Gas Pipeline (Interstate). US oil pipelines are also regulated using index-based rate caps.

---

[8] Pub. L. No. 96-448, 94 Stat. 1895 (October 14, 1980).

[9] "Report and Order and Second Further Notice of Proposed Rulemaking," FCC89-91, CC Docket No., 87-313 (April 17, 1989); and "Second Report and Order." FCC90-314 CC Docket No. 87-313 (September 19, 1990).

<u>Rate Freezes</u>

Rate freezes have been quite common in the US. Extended periods of operation without rate cases have been achieved at one time or another by many US energy utilities. These sometimes result from commitments to formal rate freezes. Freezes have often been occasioned by special circumstances such as mergers or retail competition. The FERC's rate plans for International Transmission and Michigan Transco involve rate freezes. Also noteworthy are plans for the bundled power services of AmerenUE (MO), Black Hills Power & Light (SD), Carolina Power and Light (NC), Duke Power (NC), several Michigan utilities, and Florida Power and Light (FL); for the power distribution services of National Grid (MA and NY) and Commonwealth Electric (MA); and for the gas distribution services of Consumers Energy and Michigan Consolidated Gas (MI).[10]

2.2.2 Canada

In Canada, rate indexing appears to have begun began in the telecommunications industry. The Canadian Radio-Television and Telecommunications Commission (CRTC)[11] approved a rate indexing plan for jurisdictional utilities in 1997.

The Ontario Energy Board has approved rate indexing plans for Union Gas[12] and provincial power distributors[13]. Rate indexing plan have also been approved in Canada for the power distribution services of EPCOR (ALTA) and the gas distribution services of Terasen (BC). A PBR plan with a "stair-step" mechanism has been approved for ATCO Gas North. Formal rate freezes seem to be less common in Canada than in the United States. However, Ontario power distributors recently operated under freezes for several years.

2.2.3 Britain

Rate indexing has been extensively used by regulators in Britain. It was first applied to British Telecom in 1984. Since then, rate indexing has been applied to the country's electric, gas, and water utilities.

---

[10] The plan for National Grid (MA) involves a rate freeze period as well as an indexing period.
[11] "Price Cap Regulation and Related Issues." Telecom Decision CRTC 97-9 (May 1, 1997).
[12] Ontario Energy Board. *Decision with Reasons*, RP-1999-0017, July 21, 2001.
[13] Ontario Energy Board. *Decision with Reasons*, RP-1999-0034. Jaunary 18 2000.

2.2.4 Australia

Rate indexing is also common in Australian regulation. The country's telecommunications industry has been under price controls since 1989. Rates for energy distributors in the states of Queensland and Victoria have also been subject to rate indexing.

2.3 Rate Caps and Marketing Flexibility

A major attraction of rate cap plans is the potential for enhanced utility marketing flexibility. In this section I first address the need for marketing flexibility. There follow discussions of marketing flexibility provisions under rate caps and their precedents.

The marketing flexibility provisions discussed here are particularly useful for utilities facing price-elastic demand in important markets. This situation can occur in gas delivery as well as power generation and transmission, but is comparatively rare in power distribution. Stakeholders with an immediate interest only in power distribution can save time by passing over this section.

2.3.1 Need for Marketing Flexibility

The terms on which most utilities offer their services are inconsistent with what is known about the demands for these services and the cost of providing them. Services are sometimes priced below the cost of their provision, encouraging excessive use. When capacity is fully utilized, rates may not be high enough to allocate it to users who value the services most highly. On other occasions services are priced well above the cost of provision. This discourages cost effective uses of utility services, especially in cases where demand is price elastic. Utilities also typically fail to offer the complex array of price and service options that customers' desire.

2.3.2 How Rate Caps Help

Rate caps strengthen incentives for utilities to increase the market responsiveness of their rate and service offerings. Profits can be bolstered by reducing rates in situations where they exceed cost and demand is price elastic. Utilities may also wish to use rates to discourage service requests that are unusually costly to fulfill and to encourage requests that are less costly. To the extent that they are external, price caps can also enhance the marketing flexibility that regulators can responsibly allow. That is because external rate adjustment mechanisms reduce potential concerns with cost shifting and cross-subsidization that arise when a utility's own cost and output data are used to set prices.

The amount of marketing flexibility afforded by a price cap plan depends greatly on the plan details. Two approaches are commonly used. One is automatic rate adjustments through the price cap mechanism. The other is light handed regulation of optional tariffs. We discuss each in turn.

Automatic Rate Adjustments

Price cap indexing mechanisms like those detailed in relation [1] provide a vehicle for controlled design and rebalancing of tariffed rates. The amount of automatic rate adjustment flexibility afforded by price cap plan depends in part on the specification of the actual price index. Generally speaking, an API that summarizes the escalation in several prices gives the utility some discretion in the implementation of the price escalation restrictions. In North American plans, the API is typically an explicit function of the prices of the individual services that it covers. For example, the growth in the API can be a weighted average of the growth in the prices of individual services. The weights would in this case typically be the shares of the services in total revenue.

To better understand the marketing flexibility afforded by price cap mechanisms consider, first, the case in which growth in the prices of individual services are capped but not the growth in specific *rate elements* of the services. There would in this case be a separate API for each service which summarizes the growth in the rate elements for the service. The utility can then escalate some rate elements more rapidly than the PCI so long as other elements grow less rapidly. Suppose, for example, that a utility's charge for residential energy distribution service consists of a customer (access) charge and a volumetric charge. If the PCI permitted the charge for the service to rise by 1%, the utility might then raise the customer charge by 3% and lower the volumetric charge by 1%. The price cap mechanism in this case permits automatic rate redesign.

Consider next the case in which each API summarizes the growth in the prices for a "basket" of regulated services. A utility might then raise the prices of some services by more than the PCI growth so long as rates for other services in the basket grow less rapidly. The price cap mechanism in this case permits automatic rate *rebalancing* as well as rate redesign.

Regulators often recognize the need for rate redesign and/or rebalancing but wish to control it. The price cap mechanism can provide such controls. For example, it is possible to permit the redesign of rates for individual services but not rate rebalancing. The degree of rate rebalancing can be limited by the design of service baskets. Less rebalancing is achievable to the extent that there are multiple baskets. More price elastic services might, for instance, be placed in separate baskets from less price elastic services.

Side conditions are also added to mechanisms to control the degree of marketing flexibility. A common condition is to limit the inflation in rates for a certain group of services to, say, the growth in the PCI plus 3%. Alternatively, rates for certain services may be frozen.

The approach to price cap indexing that allows the least flexibility is to limit the growth in each individual rate element of each tariff to the growth in the price cap index. In this case, individual rate elements of tariffs will typically grow at the same rate. This effectively discourages rate redesign as well as rebalancing.

Optional Rates and Services

A second common provision for marketing flexibility in rate cap plans is the ability to offer optional rates and services. These can be subject to lighter handed regulation or, in the extreme, decontrolled. Several kinds of optional offerings may reasonably be considered. One is *optional* tariffs for *standard* services. Another is new services. A third is non-essential services. A fourth is unusually complex service packages that may include standard services as components. A fifth is services to price elastic markets.[14]

Rate caps can substantially mitigate the cross-subsidy concerns that these offerings raise under COSR. That is because prices charged are not linked directly to costs and utilities have no incentive to manipulate cost allocations in a manner that creates cross subsidies.[15] By way of example, a discount offered to an economically marginal customer can affect the rates of other customers, if at all, only after the next rate case. In the meantime, the utility would only lose money if it priced its service at less than the market would bear. This encourages it to strike a

---

[14] Some services may qualify for light handed regulation under more than one of these criteria. For example, a utility might wish to offer a service that is new and inessential to a competitive market.

[15] See, *e.g.*, R. Brauetigam and J. Panzar, "Diversification Incentives Under "Price-Based" and "Cost-Based" Regulation," *RAND Journal of Economics*, Autumn 1989, 20:3, 373-391.

price that yields the greatest possible margin. Concern about cost shifting can be further mitigated by placing a floor on the optional rate that equals the incremental cost of service.

2.3.3 Marketing Flexibility Precedents

There are many precedents for marketing flexibility in rate regulation. Flexibility provisions have to date been most extensively used in the regulation of railroads, telecom utilities and oil pipelines, where the need for them is greatest. I begin with these examples to build intuition before considering precedents in energy utility industries.

Railroads

The railroad industry provides one of the most interesting experiments in marketing flexibility for a regulated industry. The need for marketing flexibility in this industry stems from both demand and supply side considerations. The demands for railroad services have varied degrees of demand elasticity. The chief source of elasticity is competition. Trucking companies, airlines, pipelines, barge lines, and lake and ocean shipping lines, as well as other railroads, may compete for cargos that a particular railroad might haul. Railroads also face indirect competition from suppliers of alternatives to the products of potential customers. Consider the case of steam coal. The demand for shipments of steam coal is sensitive to the delivered cost of natural gas, an alternative fuel, at possible generation sites. This places railroads in competition with natural gas pipelines.

Railroads also face considerable price elasticity at the margin of use. For example, an electric utility that uses coal must typically purchase both coal and coal delivery services. Purchase from more distant fields involves higher transport bills. The competitiveness of long distance shipments is thus especially sensitive to the price of transportation.

Economically marginal customers are another source of demand elasticity for railroads. That is, some customers have marginally profitable businesses and rely on the railroads for the delivery of their products or important inputs. A high cost coal mine is an example.

On the supply side, railroads must grapple with differences in the cost of requested services. For example, it is cheaper for railroads to provide service if customers ask for fewer pick up and drop off points and make fewer shipments. The distance of pick up and drop off points from major rail lines is another important cost consideration.

Policymakers have in the last thirty years recognized the marketing challenges facing railroads and afforded them extensive marketing flexibility. Confronted with an industry that provided vital services but was failing to earn its allowed rate of return, the federal governments of the US and Canada have passed a series of acts that reformed railroad regulation. In the United States, the most notable legislative initiatives have been the Railroad Revitalization and Regulatory Reform Act of 1976, the ICC Termination Act of 1995, and the Staggers Rail Act of 1980. In Canada, the Canada Transportation Act is salient. The Surface Transportation Board has promoted marketing flexibility through a series of decisions.

Consider first the US regulatory system. COSR has been largely abandoned for railroads in as much as allowed rates are rarely based on an allocated portion of a railroad's approved revenue requirement. Instead, rate restrictions, where applied, are based on estimates of stand-alone cost.

Regulation of the terms of US railroad services is limited to markets where railroads have demonstrated dominance. Services to numerous markets have been officially exempted from regulation. In other markets, simple tests are used to gauge railroad dominance.

U.S. railroads enjoy substantial marketing flexibility even where they have market dominance. For example, they are free to enter into confidential contracts with shippers. Railroads must produce formal tariffs only if a shipper requests it. The contracts commonly have tailored pricing and service quality provisions. Captive shippers can challenge the terms of service railroads offer. However, the regulations governing maximum rates to captive shippers give railroads substantial pricing discretion. Most notably, "differential pricing" is sanctioned in which rates can vary with elasticities of demand in different markets.

Marketing flexibility is also extensive in Canadian rail regulation. Under the *Canada Transportation Act*, regulation is limited to those services and regions where it is necessary to serve the needs of shippers and must not unfairly limit the ability of any carrier or mode to compete freely. As in the States, railroads can enter into confidential contracts with customers. Additionally, rates for grain shipments are subject to an indexing plan that gives railroads considerable marketing discretion.

These policy measures have made possible a fascinating experiment in how utilities facing complex and changing demands might use marketing flexibility. One striking result has been the pervasiveness of special contracts. According to one Canadian author, "confidential contracts have allowed railways and shippers to craft rate and service arrangements particular to their own needs. The concept, allowing shippers and carriers to effectively tailor their own transportation regimes, which they agree to keep confidential, has been an overwhelming success, garnering strong support from both shippers and carriers."[16]

In the United States about 70% of the tonnage of class I line haul railroads by 1997 occurred under special contracts. Another 12% of 1997 tonnage was exempt from economic regulation. Only 18% of tonnage was subject to rate reasonableness regulation.

There is abundant evidence that US railroads use the marketing flexibility they are allowed to engage in differential pricing. An example is a 1999 study by the U.S General Accounting Office (GAO) of the Carload Waybill sample that the regulator maintains.[17] The study found marked differences in the margins from services with different demand elasticities. For example, margins were greater on shipments of wheat on routes where there were few competitive transport options (*e.g.* Great Falls to Portland) than where there was competition from other railroads and other forms of transportation (*e.g.* Minneapolis to New Orleans). Low margins on motor vehicle shipments (*e.g.* Ontario to Chicago) reflect trucking industry competition.

There is also evidence that railroads adjust rates to reflect change in the markets supplied by shippers. For example, the GAO report states that rates for some shippers rise and fall with export demand.

> The volatility in commodity markets can affect railroad rates because it affects the demand for rail transportation. As demand changes, railroads adjust rates to attract or retain business. For example, officials at one Class I railroad told us that it has a wide range of pricing policies for chemicals that allow it to react to changes in world chemicals markets. Officials from the same railroad said that export demand can play a particularly strong role for grain.[18]

The western coal industry of the United States provides case study of the manner in which a transportation industry with marketing flexibility and strong marketing incentives can transform the market for shippers' products. Changes in US environmental policy have stimulated demand for the low sulfur coal that is abundant in many parts of the west. However,

---

[16] D.W. Flicker, "Canada-United States Railway Comparison: Research Conducted for the Canada Transportation Act Review, mimeo, November 2000.

[17] United States General Accounting Office, *Railroad Regulation*, *Changes in Railroad Rates and Service Quality since 1990*, GAORCED-99-93, Washington, DC, April 1999.

generating companies have other means of controlling sulfur emissions as well. These include scrubber facilities, low sulfur coal from eastern fields, and the use of gas- and nuclear-fueled generation. Western railroads have responded to this marketing challenge by offering attractive prices for long distance shipments. This has encouraged the use of western coal as far afield as Michigan and Louisiana. Railroads also use marketing efficiency to encourage shippers to use their system in cost effective manners. Western coal haulers, for instance, encourage shipment in lengthy "unit trains" devoted to particular customers.

In Canada, railroads offer rebates to grain shippers who can assemble large numbers of cars on their own.[19] Advance ordering systems are in place that offer discounts to shippers that can make advance commitments to ship certain volumes. An example is the Canadian Pacific's Max Trax plan.

Telecommunications

Incumbent telecommunications utilities ("telcos") have also faced serious marketing challenges in recent years. They have, like the railroads, faced varied degrees of competition in the major markets they serve. Markets in which competition is especially strong include those for long distance service generally and for local exchange services to larger volume customers in urban areas. Competition is much less severe in markets for small volume customers but even here there are challenges from cellular and PCS companies, cable television networks, and competitive local exchange carriers. To complicate matters further, prices to business customers have traditionally subsidized service to small volume customers (and also rural customers) in many regions.

As in the railroad industry, regulators have recognized the marketing challenges facing incumbent telcos and have granted them substantial marketing flexibility. The provision of long distance service has now been substantially decontrolled. In local exchange service, extensive use has been made of the marketing flexibility provisions discussed above.

---

[18] Ibid p. 37

[19] See James Nolan, Assessing the Impact of Bill C-34 on the Grain Handling and Transportation System in the Province of Sasketchewan.

In the United States, marketing flexibility was featured in the very first telco price cap plan, that for AT&T.[20] Service baskets were established and the growth in the API for each basket was a revenue-share weighted average of the rate elements of services in the basket. The mechanism afforded the company automatic rate redesign and rate rebalancing flexibility. However, the degree of flexibility was controlled. Separate baskets were established for residential and small business users, 800 service, and other, more competitive services. In establishing multiple baskets, the FCC explained that

> Imposing an aggregate cap on a basket of services assures regulatory control over prices charged to the class of consumers within the basket, and prevents cross-subsidization of services outside the basket by those inside.[21]

Furthermore,

> Our baskets…approach can and should be tailored to give AT&T less flexibility in its pricing of residential and various less competitive services, and greater flexibility to price efficiently in more competitive areas.[22]

Side conditions provided additional controls on the extent of rate rebalancing. For example, the average residential rate was allowed to grow by only 1% more than the price cap index each year. This general approach to marketing flexibility has since been featured in many other rate plans for US telcos.

Marketing flexibility has also been featured in Canadian telecom regulation. In the first CRTC price cap plan, all capped services were placed in a single basket. Growth in the API was a revenue-share weighted average of rate elements. Numerous side conditions were imposed to control rate rebalancing and redesign. For example, the escalation in the prices of services in two "sub-baskets" (Basic Residential Local Services and Other Capped Services) were each restricted to rise by no more than the inflation rate each year. Additionally, escalation in individual rates for residential and single-line basic services in smaller exchanges was limited to 10% each year. A few services, such as 9-1-1 service, were subject to a rate freeze.

No caps were imposed on the terms of optional services. In the CRTC's words, "Given the discretionary nature of this class of services, the Commission is of the view that an upper

---

[20] See, for example, "In the Matter of Policy and Rules Concerning Rates for Dominant Carriers", CC Docket No. 87-313 (March 1989). Regulation of AT&T rates was abandoned after long distance competition strengthened.
[21] Ibid 337, p. 166.
[22] Ibid 360, p. 180.

pricing constraint is not warranted."[23] The Commission also elected to remove from price caps certain services, such as Special Facilities Tariffs, that were "redundant or impractical" to include. However, it did not exclude services to competitive markets.

In the second CRTC price cap plan the price cap mechanism and light handed regulation were both still employed to afford telcos marketing flexibility. However, automatic rate rebalancing was further restricted by the establishment of more numerous service baskets. Side conditions were employed. For example, there were 5-10% annual caps on the escalation of individual rate elements. Rates for several services were, once again, frozen.

Optional residential services have been placed in separate baskets covered by caps. However, several services are excluded from the caps. These include, as before, business optional local services, certain complex service bundles that contain price-capped services, and certain Special Assembly Tariffs. In discussing the latter group, the Commission notes that "these services are generally offered to a limited number of customers and the rates are often developed having regard to factors such as long term customer commitments."[24] The Commission permitted telcos to offer certain services, such as Centrex, to competitive markets free from price caps.

The marketing flexibility granted to Canada's incumbent telcos brought marked changes in their rate and service offerings. Most notably, they elected to discount rates for services to larger volume customers in major metro areas substantially. Rates for residential customers, meanwhile, typically escalated by the maximum rates allowed.

Oil Pipelines

The oil pipeline industry comprises a fairly diverse set of businesses. Some pipelines transport crude oil from producing fields to refineries or storage facilities. Petroleum product pipelines transport diverse refined products (*e.g.*, gasoline, kerosene, home heating oils, jet fuels and diesel fuels) from refineries to marketing terminals.[25] Many oil pipelines are owned by shippers and some of these offer service to competitors.

An oil pipeline can face competition from other, substantially unregulated modes of transportation, as well as from other pipelines. In 2002, crude oil pipelines carried 74.7% of the

---

[23] CRTC (1997) *op cit*, 142 p. 21.
[24] CRTC (May 2002) 457.
[25] *Oil Pipelines of the United States: Progress and Outlook*, Association of Oil Pipelines, Washington, D.C.

total crude oil transported while water carries, motor carriers, and railroads accounted for 24.9%, 0.3% and 0.1%, respectively, of the total. In the same year, product pipelines carried 62.3% of the total while the other three modes carried 26.3%, 3.5% and 2.3% of the total.[26] It is plain from these statistics that water carriers are the main competitors to pipelines. However, they are able to compete with pipelines only where waterways are available.

Pipelines serve markets with varied competitive pressures. For instance, a pipeline going from a refinery in Tulsa, OK to a marketing terminal in St. Louis, MO might face competition in St. Louis from water carriers moving gasoline out of Houston. This same pipeline, however, might have more market power at locations on the way to St. Louis. Another example is a product pipeline running from the Gulf Coast along the eastern seaboard to the Northeast. It faces much less competition in Georgia and the Carolinas than it does in major cities of the Northeast, which are served by marine carriers, other pipelines, and local refineries.

The 1906 Hepburn Act brought oil pipelines under the Interstate Commerce Act, which mandated that interstate oil companies be common carriers and that they charge just and reasonable rates. From 1906 until 1977, the rates and terms of service offered by the oil pipelines were regulated by the Interstate Commerce Commission (ICC). In the initial years, the ICC was lax in regulating rates but beginning in the 1940s it applied a fixed rate of return of 8% for crude oil pipelines and 10% for petroleum product pipelines to its valuation of their assets to determine allowed revenues.

In 1977, the Department of Energy Organization Act transferred regulatory authority of oil pipelines from the ICC to the Federal Energy Regulatory Commission (FERC). The FERC initially adopted a cost based approach to regulation.[27] This decision was challenged in court and led the FERC to issue another opinion, which set out a cost-based rate methodology using trended original cost valuation.[28] In 1988, following dispute over disclosure of confidential cost information by Buckeye Pipeline, the FERC allowed a market-based rate alternative for pipelines that can show lack of significant market power.[29]

The Energy Policy Act of 1992 directed the FERC to develop a “simplified and generally applicable methodology” to regulate pipeline rates. In Order No. 561, the FERC set out a new

---

[26] *Shifts in Petroleum Transportation*, 2002, Association of Oil Pipelines, Washington, D.C.
[27] Opinion No. 154, *Williams Pipe Line Co.*, 21 FERC ¶ 61,260 (1982).
[28] Opinion No. 154-B, *Williams Pipe Line Co.*, 31 FERC ¶ 61,377 (1985).

ratemaking methodology, which uses indexing.[30] The indexing methodology caps individual pipeline rates using a price cap index that is based on an inflation measure and a productivity offset.[31] Although the indexing method freezes in place patterns of rates that existed upon its adoption, the order also permits cost-of-service proceedings that allow pipelines to request a rate above the index ceiling. In cases where pipelines can show a substantial divergence between actual cost and revenues based on rates at the ceiling level, they are allowed to charge cost-of-service rates. The FERC allows pipelines to charge market-based rates if they can demonstrate that they do not exercise significant market power in relevant markets.[32] Market-based rates allow pipelines substantial pricing flexibility in competitive markets, where rates they charge shippers fluctuate in response to changing supply and demand conditions. Order No. 561 also sets out provisions for pipelines to charge rates on negotiated basis. A version of this method that applies to existing rates, called settlement rates, allows pipelines pricing flexibility as long as they obtain "unanimous agreement" from all shippers using the rate; settlement rates can be filed that exceed the index ceiling as long as pipelines and all shippers agree on the rate. A second version of this method, which applies to new rates and is simply called negotiated rates, requires a pipeline to secure an agreement with a non-affiliated shipper to file this rate offering. Both methods allow pricing flexibility for pipelines as long as they do not use market power to 'coerce' agreement from shippers.

These policy measures give pipeline companies substantial flexibility to respond to competition and to develop tailored service packages for customers. Market-based or negotiated rates allow pipelines to meet competition and take advantage of business opportunities. Indexation protects customers in less competitive markets and provides a potentially useful means of updating the terms of special contracts.

---

[29] *Buckeye Pipe Line Co.*, 44 FERC ¶ 61,066 (1988).

[30] Order No. 561, *Revisions to Oil Pipeline Regulation Pursuant to Energy Policy Act of 1992*, III FERC Stats. & Regs. ¶ 30,985 (1993).

[31] Following a review of the indexing rate, as required by Order No. 561, the Commission issued a December 2000 order affirming the method. *Five-Year Review of Oil Pipeline Pricing Index*, 93 FERC ¶ 61,266, Docket No. RM00-11-000 (December 14, 2000). The Association of Oil Pipelines challenged this order in court and upon further review the FERC limited the index to track an inflation measure only. *Five-Year Review of Oil Pipeline Pricing Index: Order on Remand*, 102 FERC ¶ 61,195, Docket Nos. RM00-11-000 and RM00-11-001 (February 24, 2003).

[32] The FERC issued Order No. 572, which details the requirements for application of market-based rates. It also indicates that pipelines can not charge rates above the index ceiling until the Commission finds they lack significant market power. Order No. 572, *Market-Based Ratemaking for Oil Pipelines*, III FERC Stats. & Regs. ¶ 31,007 (1994).

Gas and Electric Utilities

The precedents for marketing flexibility are not as extensive in the gas and electric utility industries. This reflects, in part, less acute competitive challenges. However, all of the major marketing flexibility provisions discussed above do have precedent. The use of a price cap mechanism to permit automatic rate redesign and rebalancing, for instance, has been approved in a number of jurisdictions. Most notably, the Netherlands Electricity Regulation Service (DTE) has approved this approach for the regulation of TenneT, the Dutch power transmission monopoly, and for Dutch gas and electricity distributors. In the case of TenneT, the growth rate in the API for network service is a weighted average of the growth rates in a standing charge, a volumetric charge, an annual capacity payment, a monthly capacity payment, and a reactive power charge.[33]

In North America, a price cap plan for Boston Gas permitted automatic rebalancing of gas distribution rates to reduce interclass subsidies and increase price signal efficiency. The Company was prohibited, however, from pricing services below marginal cost. Automatic rate rebalancing was proposed by the OEB staff for provincial power distributors in the first draft *Rate Handbook*. The proposal was rejected by the Board in its final decision[34].

Light-handed regulation of optional rate and service offerings has considerable precedent in North American energy regulation. The California Public Service Commission has allowed Southern California Gas to offer negotiated rates and optional tariffs provided the price is not less than the long-run marginal cost. In Ontario, the OEB has approved light handed regulation for certain services of Union Gas.

In 1996, the FERC issued a policy statement supporting expedited approval of negotiated gas transmission services provided that customers continued to have recourse to a rate that is based on cost of service principles. It also concluded that “where a natural gas company can establish that it lacks significant market power, market-based rates are a viable option for achieving the flexibility and added efficiency required by the current marketplace.”[35] In discussions before the FERC, pipelines have cited the need for flexibility to address a number of marketing challenges, including competition from other pipelines, the dual-fuel capability of

---

[33] Network service excludes system support services and generator connections.
[34] Ontario Energy Board. *Decision with Reasons*, RP-1999-0034, Jaunary 18 2000.
[35] 74 FERC 61,076 (January 1996) p. 8.

many large volume customers, the existence of a secondary market for firm capacity, and the desire of new electric generators for price certainty.

In the electric power industry, pricing flexibility was featured in rate plans for two Maine electric utilities, Central Maine Power (CMP) and Bangor Hydro-Electric (BHE) in the mid-1990s. Both companies were bundled power service providers at the time and had high operating costs and a number of economically marginal and/or price sensitive large volume customers. Special contracts with customers had previously been subject to lengthy investigations. The Maine Public Service Commission would approve them if it determined that the customer would not have remained a customer at the tariffed rate and the discount agreed to was not larger than necessary to keep the business.

A change in state law expressly permitted the Commission to authorize pricing flexibility programs where companies could discount rates with more limited Commission oversight. A price cap plan was approved for CMP that gave it flexibility to discount rates for standard services, develop new customer classes for targeted services, and to enter into special rate contracts with individual customers without Commission approval. All offerings were subject to marginal cost floors. The Maine Public Utilities Commission, in approving the plan, stated that

> Captive customers are protected by the rate cap and revenue deficits borne by shareholders... Because CMP will have substantial exposure to revenue losses due to discounting, the Company will have strong incentive to avoid giving unnecessary discounts, and it will have a strong incentive to find cost savings to offset any such losses. Pricing flexibility gives CMP the opportunity to use price to compete to retain customers. These features of the ... pricing flexibility program simulate conditions in competitive markets and will help the Company adapt to increasing competition in its industry.[36]

Similar language appeared in the Order approving BHE's plan. The Maine Commission has since approved pricing flexibility for power distributors in the state who operate under price caps using similar reasoning.[37]

CMP and BHE used the marketing flexibility granted under the first plan to offer special discounts to customers. This created the issue of who was to absorb the lost margin (called

[36] Re Central Maine Power Company, Docket 92-345 (II), January 10, 1995, p. 24.

[37] See, for example, *Central Maine Power Company: Annual Price Change Pursuant to the Alternative Rate Plan,* Docket No. 99-155, 13 July 1999

"revenue delta") from discounts at the time of the next rate case: the companies or other customers. In addressing this issue for BHE the Maine Commission stated that

> We remain convinced that pricing flexibility decisions should not be treated like ordinary utility expenditures in which prudence investigations provide the insurance that utility actions have been reasonable. The best means to protect ratepayers from unreasonable price discounts is to adopt an incentive mechanism like a price cap in which future rate increases are unrelated to the amount of discounts granted. It is simply too difficult and expensive to realistically review the utility's actions and customers' alternatives that resulted in the utility's granting a price discount.[38]

After considering the riskiness of the unnecessary price discounts that BHE faced under its first plan, the Commission decided to allocate the lost margins 85% to ratepayers and 15% to shareholders.

In Ontario, where Energy Board Staff recommended pricing flexibility for provincial power distributors in its proposed electric distribution *Rates Handbook*,[39] staff stated in the draft *Handbook* that

> One overall price cap for a utility that imposes an average adjustment to all prices may prove unsatisfactory from several perspectives including limiting a distributor's ability to fine tune its cost allocation and its responsiveness to pricing pressures in particular sub-markets.[40]

Staff also cited the usefulness of pricing flexibility in achieving gradual rate harmonization after mergers. Staff provided an example of automatic rate rebalancing through a price cap mechanism. The pace of rebalancing would be controlled via side conditions. The Board rejected Staff's proposal in its *Rates Handbook* decision.[41]

2.4 Evaluation of Rate Caps

Rate caps can generate utility performance incentives much stronger than those obtained under typical cost of service regulation. One reason is that incentives are comprehensive so that a wide range of cost containment, product development, and marketing initiatives are encouraged. Another is that indexing can facilitate an extension of the period between rate cases. To the extent that this is true, improved unit cost performance does not reduce allowed

---

[38] *Bangor Hydro Electric: Proposed Increase in Rates,* Docket No. 97-116, March 24, 1998.
[39] Ontario Energy Board Staff, *Proposed Electric Distribution Rates Handbook,* mimeo, June 30, 1999.
[40] *ibid*, p. 4-9.
[41] Ontario Energy Board. *Decision with Reasons*, RP-1999-0034, January 18 2000.

price escalation during the term of the plan. The benefits of improved performance can thus go straight to the bottom line. The potential impact on productive and allocative efficiency is substantial. The actual incentive effects of rate caps depend greatly on plan details. For example, incentives increase with the length of the indexing period and with the introduction of post plan sharing provisions.

Rate caps can provide a further boost to efficiency by permitting a relaxation of operating restrictions. The case of marketing flexibility is illustrative. To the extent that rate restrictions are external, customers of monopoly services can be insulated from the effects of a company's operations in competitive markets. This reduces concerns about cross subsidization. Light-handed regulation of utility rates for non-core services is then possible. A company can also have more leeway in its purchases from affiliates and its depreciation practices.

Rate caps can reduce regulatory cost. Some startup costs must, of course, be incurred to master the new regulatory system. These may include a close monitoring of the company's operations during the terms of the first indexing plans. But the frequency of future rate cases can be substantially reduced. Furthermore, reliance on external indexes diffuses inherently controversial cost allocation and transfer pricing issues. On the other hand, controversy can be considerable over alternative methods for measuring input price and productivity growth.

The numerous inherent advantages of rate caps are offset to some degree by disadvantages. One is regulatory risk. The novelty of rate indexing encourages the selection of key plan parameters arbitrarily. Utilities may reasonably worry that that regulators will choose plan terms that prevent the recovery of prudently incurred cost. Customers may reasonably worry that plan terms will deny them a fair share of plan benefits. Concerns about arbitrary selection of key plan parameters reduce the willingness of parties to try the rate-indexing option and can weaken the incentive benefits of price cap plans substantially. A rate freeze is a sensible alternative to indexing in jurisdictions where this is a concern but is not suitable in all times and places, as has been noted.

Rate caps also involve business risk: the possibility that price restrictions will not track trends in external business conditions that affect a company's unit cost. Relevant business conditions include weather, the business cycle, input prices, and government policy. Windfall gains and losses may occur if rate caps don't reflect changes in these conditions.

3 REVENUE CAPS

Under a revenue cap, the revenue of a utility is the focus of control. Two approaches to escalation of a comprehensive revenue requirement escalation are noteworthy. One is comprehensive indexing. Another is the disaggregated approach in which there are separate escalation provisions for O&M expenses and capital.

3.1 Comprehensive Revenue Caps

3.1.1 Comprehensive Indexing

<u>The Basic Idea</u>

Under this approach, the growth of the revenue requirement is usually limited to the growth in a revenue cap index (*RCI*), as in the following formula:

$$\textit{growth Revenue Requirement} \leqq \textit{growth RCI} \qquad [2]$$

Like *PCIs*, *RCIs* often feature measures of price inflation. *RCIs* may include, additionally, a measure of output growth.[42]

The addition of a balancing account mechanism can ensure that actual revenues are similar or equal to the revenue requirement. The balancing account contains the value of any mismatch between actual revenue and the revenue requirement until rates can be adjusted to eliminate it. These arrangements are sometimes called revenue-decoupling mechanisms since they sever the link between revenue and efforts to market regulated services.[43]

Revenue cap mechanisms typically do not specify how revenue limits are translated into rate limits. Service offerings and the fashioning of rates from revenue can, in fact, continue using traditional methods. The utility can, in principle, be afforded some flexibility in the provision of rate and service options. However, incentives for efficient marketing are weaker than under a rate cap mechanism, as I discuss further below.

---

[42] This is discussed further in Section 5.

[43] Revenue decoupling mechanisms have also been used in the absence of indexing. Prominent examples include the electric revenue adjustment mechanisms that have been used in California and Maine.

Precedents

United States Southern California Gas has operated under comprehensive revenue caps since 1996. Comprehensive revenue caps have also been approved for the power distribution services of Pacific Gas & Electric, Pacificorp, SDG&E, and Southern California Edison (SCE) and for the gas distribution services of Baltimore Gas & Electric (“BG&E”), PG&E, SDG&E.

Canada The National Energy Board (NEB) of Canada has approved comprehensive revenue caps for two oil pipelines, Enbridge Pipelines (formerly Interprovincial Pipe Line) and TransMountain Pipe Line.

*Britain_* The power transmission services of National Grid have been subject to revenue caps since 1993. All regulated transmission services were subject to revenue caps under the first plan. Dispatching and other system operation services have since been exempted from revenue caps.

Australia Revenue caps are used by the ACCC to regulate power transmission services of Energy Australia, Powerlink Queensland, Powernet Victoria, and Trans Grid in Australia. The inflation factors in all of these plans are consumer price indexes. Comprehensive revenue caps are also used to regulate power distributors in New South Wales.

3.1.2 Disaggregated Approach

The Basic Idea

Under the disaggregated approach to comprehensive revenue caps, components of the revenue requirement that correspond to different costs are accorded different ratemaking treatment. A distinction is commonly drawn between O&M expenses and capital. A salient advantage of this approach is the ability to tailor the capital cost allowance to the special circumstances of a utility.

Precedents

United States California was a noteworthy early practitioner of the disaggregated approach to comprehensive revenue requirement escalation. Since the 1980s, COSR in California has taken the form of a general rate case (GRC) cycle in which rate cases were typically held every three years and a mechanism adjusted the revenue requirement more or less automatically in the two out years.[44] O&M expenses were subject to indexation whereas capital cost was subject to a more cost of service treatment where, for example, capital additions over a multi-year period were assumed to equal the actual additions over a recent three year period. In many plans, the cost of funds was subject to adjustment each year in a streamlined "MICAM" proceeding.

Innovation to the California attrition formula occurred in the first PBR plan for San Diego Gas and Electric. This plan, which applied to both gas and electric services, was approved in 1994.45 It featured separate index-based adjustments for revenue requirements corresponding to allowed O&M expenses and capital spending. Separate O&M indexing mechanisms were specified for gas and electric operations. The mechanisms included inflation factors, X-factors, and adjustments for output growth.[46]

Canada The disaggregate approach to revenue caps has been more widely used in Canada than in the US. BC Gas began operating under caps for certain categories of base rate revenue in 1994. The caps pertained to O&M expenses and small capital expenditures. BC Gas also operates under a revenue decoupling mechanism called the Revenue Stabilization Adjustment Mechanism. It applies only to revenues from residential and commercial sales.

The NEB approved a disaggregated revenue cap mechanism for gas transmission services of Westcoast Energy in 1996. Indexing limited growth in the revenue requirement components covering O&M expenses and small capital additions. The formula for growth in both revenue cap indexes was forecasted inflation in a CPI. There was no explicit X or output factors in the formula.

The Alberta commission has approved disaggregated revenue caps for NOVA Gas Transmission. The caps apply to O&M expenses and small capital additions. A plan was approved by the OEB for the gas delivery O&M expenses of Toronto-based Consumers Gas in 1998.

---

[44] Such mechanisms were previously called "attrition" mechanisms but are currently called "post-test year" mechanisms.

[45] It has been claimed that the term "performance based ratemaking" was coined by San Diego personnel during this plan's development.

[46] This plan was succeeded by the rate cap plan that is mentioned above.

3.1.3 Evaluation of Comprehensive Revenue Caps

Comprehensive revenue caps can create strong incentives for cost containment by permitting operation for an extended period with an externalized revenue requirement. There are incentives for a wide range of cost containment initiatives. The external basis for the cap also encourages some forms of operating flexibility. For example, extended utility operation under a revenue cap could permit a regulator to relax restrictions on purchases from affiliates.

One important difference between the consequences of rate and revenue indexing lies in the marketing of utility services. Incentives for improved marketing are generally weaker than under rate caps. Marketing incentives may, in fact, be weaker than under COSR. For example, reducing rates for services in price elastic applications may, by raising total revenue, lower rates promptly. Utilities may, as a consequence, be less aggressive in promoting system use, including efforts to avoid uneconomic bypass. They do, however, have an incentive to raise rates when rates fail to cover the incremental cost of service.

Revenue caps can raise more concerns than rate caps about the quality of core services. As with rate caps, quality may suffer since there are strong incentives to cut costs. While the pressures to minimize costs are similar under rate and revenue caps, under a revenue cap revenues that are lost if poor service leads to fewer sales can be recovered through price increases on remaining customers using the balancing account. Since this is not possible under rate caps, the incentives to maintain service quality are weaker in the absence of counterbalancing incentive provisions. This concern will be greater to the extent that customers care about quality and lack cost-competitive alternatives.

Revenue cap plans reduce windfall gains and losses from demand fluctuations. This is an important consideration for utilities that face unusually volatile demand due, for instance, to sensitivity to weather, prices of competing products, or prices in the end product markets of business users. Stabilization of revenue can lower a utility’s capital cost but in the process destabilizes rates. For example, a recession in the service territory can place upward pressure on rates at times when rate increases are especially unwelcome. Plan designers thus encounter the issue of whether the benefits of capital cost savings to customers offset the cost of greater rate stabilization.

Another important attribute of revenue caps is their ability to strengthen the incentives to promote energy conservation. Under rate caps, the promotion of conservation can reduce a utility's operating margins. Under revenue caps, rates rise automatically to offset this effect[47].

Consideration should also be given to the issue of regulatory cost. Revenue caps can permit economies in the cost of regulation relative to COSR. However, regulatory cost is likely to be somewhat greater than under rate caps. One reason is the need for periodic filings to implement the balancing account mechanism. There may, additionally, be a continued need to consider the allocation of revenue requirements between customer groups, service offerings, and rate design. Note that the addition to the indexing formula of an output growth factor creates another potential plan design controversy.

3.2 Non-Comprehensive Revenue Caps

3.2.1 The Basic Idea

Under non-comprehensive revenue caps there are caps on only a portion of the company's revenue requirement. An example might be a cap on the revenue requirement (allowed cost) for O&M expenses. Partial revenue caps are, like comprehensive caps, often fashioned using indexes. In an indexing mechanism, an adjustment for output quantity growth is once again needed. As with comprehensive revenue cap plans, partial indexing plans typically do not address rate and service offerings. Utilities therefore typically require authority outside of partial rates and revenue caps to alter these offerings.

3.2.2 Precedents

A plan for the allowed O&M expenses of Consumers Gas in Ontario[48] is a good example of a non-comprehensive revenue cap. Allowed expenses were escalated by an index featuring an inflation measure and adjustments for price and productivity growth.

[47] The use of revenue caps for this express purpose has been likened, however, to sending an elephant to do the job of a mouse since DSM typically has a comparatively small impact on volumes compared to other demand drivers, such as weather.

[48] Ontario Energy Board. *Decision with Reasons*, E.B.R.O. 497-01, April 22, 1999.

3.2.3 Evaluation

Non-comprehensive revenue caps can make revenues in the targeted areas less sensitive to the operations of the subject utility. This can substantially strengthen incentives to contain the associated costs. It can also permit increased operating flexibility in the targeted areas. Suppose, by way of example, that a utility wishes to purchase many of its O&M services from unregulated affiliates. A cap on allowed O&M expenses can then permit relaxed vigilance on service transfers without placing recovery of capital cost at risk. Of course, a *comprehensive* rate or revenue cap plan would also accomplish this. The approach can focus management attention on specific problems and help accelerate their rectification. A partial indexing approach is also useful where there is consensus only to apply PBR to certain areas of the company's business. If the utility is undergoing pro-competitive restructuring, for instance, plans may be designed to focus only on areas subject to continuing regulation.

Non-comprehensive revenue caps can make sense in situations where comprehensive caps don't make sense. One example is a situation where company expects to make a sequence of large capital additions in the next few years. The company has a legitimate concern about the recovery of these costs, and may wish for this reason to see them approved and included in the rate base. On the other hand, a sequence of traditional rate cases will weaken incentives for O&M cost management.

One potential problem with partial revenue caps is the unevenness of performance incentives that result. There may be less incentive to control cost in non-targeted areas. The company may, in the extreme, be given an incentive to improve performance in the targeted areas *at the expense of* performance in other areas. If a utility were subject only to a cap on O&M revenue, for instance, excessive capital spending could be undertaken to reduce O&M expenses. Overall, the company's performance might not improve. This problem is mitigated to the extent that the partial caps cover most areas of controllable cost. For example, plans covering both O&M expenditures and capital expenditures can be defended on the grounds that they cover all "controllable" costs.

Partial revenue caps share with comprehensive caps several other attributes. One is relatively weak incentives for better marketing. Another is stronger incentives to promote energy conservation.

4 INDEX DESIGN ISSUES

Rate and revenue cap indexes can materially extend the period between rate cases by providing automatic adjustments for input price inflation and other changes in business conditions that affect the unit cost trends of utilities. The design of such indexes is therefore a salient issue in many PBR proceedings. In this section I discuss key capping index design issues.

4.1 Overview

4.1.1 Index Formulas

Price cap index formulas vary from plan to plan but have the following general structure. The PCI growth rate $(growthPCI)$ is the difference between an inflation factor (*P*) and an X-factor (*X*), plus or minus a Z-factor (*Z*),[49] as in the following formula:

$$growth\ PCI = P - X \pm Z. \qquad [3]$$

Compared with price cap indexes, a growth rate formula for a revenue cap index requires an additional adjustment to reflect the effect of output growth on cost. Some approved RCI formulas have an explicit term for such an adjustment that may be called an output factor and is here denoted by *Y*.

$$growth\ RCI = P - X + Y \pm Z\ . \qquad [4]$$

The X and Y terms as here described are sometimes captured in a consolidated *X* so that the growth rate formula resembles that for a price cap. The X factor would in this case be lower by an amount that reflects the expected annual pace of output growth (*e.g.* 1-2%). If the unconsolidated *X* happens to be similar to the expected growth of output (*i.e.*,$Y = X$), the X factor can be eliminated, as in the following formula:

$$growth\ RCI = P \pm Z\ . \qquad [5]$$

If, alternatively, the inflation rate is deemed to be similar to the appropriate *X*, the rate cap growth formula can reduce to

$$growth\ RCI = Y \pm Z\ . \qquad [6]$$

---

[49] The term Z-factor appears to have developed in the FCC proceeding to develop a price cap plan for AT&T. It was so called because the PCI for AT&T also included an X-factor as here described and a "Y" factor to affect a specific category of price cap adjustments.

Some revenue cap index plans restrict growth in revenue *per customer*. This can be shown to be mathematically equivalent to revenue requirement indexing where the number of customers is the output factor.

4.1.2 Inflation Measures

The inflation factor, *P*, provides an automatic adjustment to the PCI for price inflation. It is sometimes fixed in advance but is more commonly updated annually to reflect the recent growth rate in an external price inflation measure. Three basic kinds of inflation measures have been used in approved indexing mechanisms. We consider these in turn.

<u>Macroeconomic Measures</u>

The Basic Idea Macroeconomic inflation measures are summary measures of growth in the prices of a wide range of the economy's goods and services. Those used in PBR plans are typically output price indexes computed by government agencies.[50] Examples include price indexes for gross domestic product (GDP-PIs) and consumer price indexes (CPIs).

In the United States and Canada, the GDP-PI is the federal government's featured index of inflation in the domestic economy's final goods and services. It differs from the CPI chiefly in covering inflation in the prices of capital equipment used by industry as well as inflation in consumer product prices. The GDP-PI is generally favored over the CPI in North American rate escalation indexes, for several reasons. Its broader coverage makes it more stable and more reflective of inflation in the prices of base rate inputs than is the CPI, which places a heavier weight on price-volatile energy and food products. An additional advantage is that the correspondence of the GDP-PI to the entire economy makes it easier to adjust the escalation formula for a possible mismatch between industry-specific and macroeconomic price inflation.

Precedents Macroeconomic inflation measures are almost universally used in telecom utilities' rate-cap plans. For example, the GDP-PI has been employed in both price cap plans of the CRTC. Macroeconomic inflation measures have also been employed in the majority of indexing plans for energy utilities. The price cap index for Union Gas, for instance, used as an inflation

---

[50] The Federal governments of the United States and Canada also produce macroeconomic indexes of inflation in the prices of several kinds of inputs (*e.g.*, labor and producer goods). These have rarely been used as stand-alone inflation measures in PCI construction due in part to the fact that each index covers only some of the relevant inputs. A prominent exception has been the use of a producer price index (PPI) in the indexing plan for US oil pipelines.

measure the GDP-PI. Consumer price indexes such as Britain's retail price index (RPI) are used in almost all overseas indexing plans.

Appraisal One advantage of macroeconomic inflation measures is their simplicity. Another is their credibility, since they are typically computed with some care by government agencies. Still another is their familiarity to stakeholders in the regulatory process. The main concern with macroeconomic inflation measures is their ability to track growth in the prices of utility inputs. This is discussed in further detail below.

Industry-Specific Measures

The Basic Idea Industry-specific input price indexes are expressly designed to track inflation in the prices of the relevant utility inputs. Such measures summarize the growth in subindexes that are chosen to track trends in the prices of major input categories. The index formula customarily assigns weights to the subindex growth rates that reflect the shares of the input categories in utility cost. This is the approach to index weighting that best captures the impact of growth in various input prices on cost.

Precedents An industry-specific input price index was first used in the PBR plan, mentioned above, for US railroads. It was called the Index of Railroad Cost. The growth rate of the index --- called the rail cost adjustment factor --- was a weighted average of the growth rates in indexes of the prices of railroad inputs. The input categories included labor, fuel, materials, equipment rentals, depreciation, interest, and a miscellaneous input category. Each input was assigned a weight that reflected its share of the cost of railroad operations nationwide.

An energy industry-specific input price index was first approved for the bundled power services of PacifiCorp (CA). The staff of the California Public Service Commission (CPUC) played an instrumental role in the index design. Industry-specific inflation measures have since been approved for the gas delivery services of Southern California Gas (CA), the gas and electric power delivery services of San Diego Gas and Electric (CA), and the power distribution services of Ontario utilities. The index approved by the Ontario Energy Board was called an industry price index (IPI).

The price inflation index in SDG&E's price cap plan for power distribution is a good example of the genre. The growth rate in this measure was a weighted average of the growth rates in price subindexes for seven input categories: capital, labor, non-labor O&M inputs used in distribution, customer accounts, customer service and information, sales, and administrative and general functions. The subindex for labor was the average hourly earnings of workers nationwide in electric, gas, and sanitary services. This was collected by the Bureau of Labor Statistics ("BLS") of the US Department of Labor. The five subindexes for non-labor O&M inputs were also of national character and were calculated by a firm that is now called Global Insight.

Pros and Cons By design, an industry-specific input price index tracks industry input price fluctuations better than an economy-wide measure. This advantage is important to the extent that the input price growth of a utility industry differs from that of the economy. For example, energy transmission and distribution are unusually capital intensive businesses and are therefore unusually sensitive to change in the cost of funds. This has a pattern of fluctuation that can differ from that of other utility inputs for extended periods. The construction of appropriate indexes is aided by well-established theory and publicly-available data.

One disadvantage of the industry-specific approach is that no official source computes input price indexes for energy utilities. Since accurate price indexes are complex, it is left to the regulatory community to hash out the best possible design. The design of the capital price index is especially complex and controversial.

An interesting issue in considering industry-specific inflation measures is their effect on regulatory risk. These measures can in principle reduce operating risk and help sidestep controversy over adjustments that may be needed to a PCI with a macroeconomic inflation measure to help it better track industry input price trends. On the other hand, approved industry-specific measures may not do the best possible job of tracking industry input price inflation. A good example is the measure approved in Ontario for power distribution. This counted only half of the calculated growth in the capital price in the name of rate stabilization. Since the capital price had the largest weight in the growth rate formula, the effect on allowed price escalation was substantial. This could matter greatly in the long run given the capital intensiveness of the distribution business.

Output Price Measures

Industry specific *output* price indexes are indexes of the prices charged for the service in question by other providers. Such indexes can potentially reflect the unit cost trend of the relevant industry. They therefore reflect industry-specific input price and productivity trends yet manage to sidestep controversial issues such as the best measure of capital price inflation.

An early example of peer price indexing was their use in plans for the rates of two bundled electric service providers in the Midwestern US. These plans linked industrial rates for each utility to rates for similar services of other utilities in the region. In North America, it has until recently been difficult to regulate most transmission and distribution services using peer price indexes due to the lack of unbundled price data on the services. However, the restructuring of the industry is now sufficiently advanced to make peer price indexing possible in the United States if not Canada. The current PBR plan for National Grid USA in Massachusetts escalates its power distribution rates using an index of the escalation in the rates charged for the same service by electric utilities in the Northeast US. The Massachusetts DTE recently approved the first escalation under the plan: 3.5%.

4.1.3 X Factors

The X-factor term of a rate escalation index is an external parameter that typically causes the index to grow more slowly than the inflation measure, to the benefit of customers. Unless the X factor is negative, prices for regulated services will fall in "real" terms. X is sometimes called a "productivity factor" or a "productivity offset" to the inflation measure since considerations of productivity growth have a bearing on its value.

Various methods have been used to ensure the external character of X. Most commonly, its value in each plan year is set in advance and is constant throughout the plan. However, in several approved plans the X-factors are set in advance but scheduled to vary from year to year. For example, X-factors in several cases have been scheduled to rise gradually over the term of the plan.

Another issue in X factor design is whether to set its value(s) in advance or, instead, to make periodic updates during the plan to reflect newly available information about business conditions. The best known precedent for this approach is the X-factor in the price cap index for

US railroads.[51] This was an annually updated rolling average of the recent productivity growth of the railroad industry. Each railroad subject to the plan affected the resultant X factor but not by enough to materially slacken its performance incentives.

4.1.4 Z-Factors

The Z-factor term of a price cap index adjusts the allowed rate of price escalation for external developments that are not reflected in the inflation and X-factors. It is apt to differ from period to period. One of the primary rationales for Z-factor adjustments is the need to adjust price caps for the effect of changes in tax rates and other government policies on the company's unit cost. Absent such adjustments, policymakers can adopt new policies that increase the company's unit cost, confident in the knowledge that its earnings, rather than its rates, will be affected. Another rationale for Z-factors is to adjust for the effect of miscellaneous other external developments on industry unit costs that are not captured by the inflation and X-factors. Z-factors can potentially reduce operating risk and discourage opportunistic behavior without weakening performance incentives since they are triggered only by external developments. Disadvantages include the fact that they can significantly raise regulatory cost, and the possibility that they may weaken utility incentives to mitigate the impacts of triggering events.

Most index-based price cap plans have explicit rules for Z-factor adjustments. Those approved by the OEB for provincial power distributors and recorded in Chapter 5 of the March 9, 2000 distribution rates handbook are illustrative:

- Causation – the expense must be clearly outside of the base upon which rates were derived.
- Materiality – the cost must have a significant influence on the operation of the utility, otherwise they should be expensed in the normal course and addressed through organizational productivity improvements.
- Inability of Management to Control – to qualify for Z-factor treatment, the cost must be attributable to some event outside of management's ability to control.

Examples include a tax change or requirements of the IMO that result in expenditures by the distribution utility. On the other hand, an ice storm that causes extensive damage in a system with sub-par maintenance would not qualify for Z-factor treatment.

---

[51] This is discussed in more detail below.

- Prudence – the expense must have been prudently incurred. This means that the option selected must represent the most cost-effective option (not necessarily least initial cost) for ratepayers.

For example, some utilities will need to upgrade their billing systems to deal with market opening. The prudence standard requires that the utility justify purchasing a new system versus outsourcing the function to a vendor, association, or utility.

While there is no general language about the relevance of government policy changes, the explicit mention of tax changes and Independent Market Operator requirements in the *Rate Handbook* is noteworthy. The Board shed further light on events it deems relevant for Z factoring in its Union Gas decision. It states that

> The Board agrees with the intervenors that the use of Z-factors limited to changes in legislative and regulatory requirements and generally accepted accounting principles specific to the natural gas business is appropriate.[52]

Regarding other kinds of events, the Board rejects a request by Union Gas to "pre-approve" stranded costs that might arise during the plan period. It does, however, state that "the Company is free to bring before the customer review process any proposals related to the recovery of stranded costs."

4.2 Index Design Methods

Two general approaches to the design of rate and revenue cap indexes have now been established: the North American approach and the British. These are so-named because of their region of origin.

4.2.1 The North American Approach to Index Design

Although index-based regulation is associated in the minds of many with Great Britain, North America actually has a longer history with this regulatory system. E. Fred Sudit of Rutgers University outlined the approach to PCI design that has become common in North America in a 1979 paper.[53] William Baumol, then at Princeton University, elaborated on the

---

[52] Ontario Energy Board. *Decision with Reasons*, RP-1999-0017, July 21, 2001.

[53] E. Fred Sudit, "Automatic Rate Adjustments Based on Total Factor Productivity Performance in Public Utility Regulation," in *Problems in Public Utility Economics and Regulation* ed. M. Crew, Lexington Books, 1979.

idea in a 1982 paper.[54] While these early treatises influenced the American approach to PCI design, credit must also go to other individuals, including several involved in the early regulatory proceedings and supporting legislation. The history of regulation using indices will be in a later section.

Logic of Price Cap Indexes

The North American approach to index design is founded on the logic of economic indexes. The analysis begins with consideration of the growth in the prices charged by an industry that earns, in the long run, a competitive rate of return. In such an industry, the long-run trend in revenue equals the long-run trend in cost.

$$trend\ Revenue = trend\ Cost. \qquad [7]$$

The assumption of a competitive rate of return is applicable to utility industries and even to individual utilities. It is also applicable to unregulated and competitively structured markets. Consider, now, that the trend in the revenue of any firm or industry is the sum of the trends in appropriately specified output price and quantity indexes.

$$trend\ Revenue = trend\ Output\ Quantities + trend\ Output\ Prices\,. \qquad [8]$$

Relations [7] and [8] together imply that the trend in an index of the prices charged by an industry earning a competitive rate of return equals the trend in its unit cost index.

$$trend\ Output\ Prices = trend\ Cost - trend\ Output\ Quantities = trend\ Unit\ Cost\,. \qquad [9]$$

The long run character of this important result merits emphasis. Fluctuations in input prices, demand and other external business conditions will cause earnings to fluctuate absent adjustments in production capacity. Fluctuations in certain expenditures that are made periodically can also have this effect. An example would be a major program of replacement investment for a distribution system with extensive asset depreciation. .

The result in [9] provides a conceptual framework for the design of a rate adjustment index that we call the industry unit cost paradigm. Suppose for example, that the growth in a utility's rates is measured by an actual price index and that a PCI limits the growth in this index. A stretch factor established in advance of plan operation can be added to the PCI growth formula

---

[54] William J. Baumol, "Productivity Incentive Clauses and Rate Adjustment for Inflation," *Public Utilities Fortnightly*, July 22, 1982, pp. 11-18.

which slows growth to the benefit of customers.[55] A PCI is then *calibrated* to track the industry unit cost trend to the extent that

$$\textit{trend PCI} = \textit{trend Unit Cost Industry - Stretch Factor}. \qquad [10]$$

A properly designed PCI that is calibrated in this manner provides automatic adjustments for trends in external business conditions that affect the unit cost of utility operation. It can therefore reduce utility operating risk without weakening performance incentives. This constitutes a remarkable advance in the technology for utility regulation.

The design of PCIs that track the industry unit cost trend is aided by an additional result of index logic. It can be shown that the trend in an industry's *total* cost is the sum of the trends in appropriately specified industry input price and quantity indexes.

$$\textit{trend Cost} = \textit{trend Input Prices} + \textit{trend Input Quantities}. \qquad [11]$$

It follows that the trend in an industry's *unit* cost is the difference between the trends in industry input price and TFP indexes.[56]

$$\textit{trend Unit Cost} = \textit{trend Input Prices} - \textit{trend TFP}. \qquad [12]$$

Furthermore, a PCI can be calibrated to track the industry unit cost trend if it satisfies the following formula:

$$\textit{trend PCI} = \textit{trend Input Prices Industry} - \left(\textit{trend TFP Industry} + \textit{Stretch Factor}\right). \qquad [13]$$

An important issue in the development of a PCI is whether it should be designed to track short run or long run unit cost growth. An index designed to track short run growth will also track the long run growth trend if it is used over many years. The alternative is to design the index to track only long run trends. Different approaches can, in principle, be taken for the input price and productivity components of the index.

One issue to consider when making the choice is the manner in which short-run input price and productivity fluctuations affect the prices charged by unregulated industries. Inflation in the prices charged by such industries sometimes accelerates (decelerates) rather promptly

---

[55] Mention here of the stretch factor option is not meant to imply that a positive stretch factor is warranted in all cases.

[56] Here is the full logic behind this result:

$$\begin{aligned}
\textit{trend Unit Cost} &= \textit{trend Cost - trend Output Quantities} \\
&= \left(\textit{trend Input Prices} + \textit{trend Input Quantities}\right) - \textit{trend Output Quantities} \\
&= \textit{trend Input Prices} \\
&\quad - \left(\textit{trend Output Quantities - trend Input Quantities}\right) \\
&= \textit{trend Input Prices} - \textit{trend TFP}
\end{aligned}$$

when input prices accelerate (decelerate). Airlines and trucking companies, for instance, sometimes hike prices in periods of rapid engine fuel price inflation.

An analogous result does not obtain for TFP since the prices charged by unregulated industries do not necessarily fall in the short-term when TFP falls. For example, TFP typically falls (rises) in the short run in response to a slackening (strengthening) of demand. These same developments typically have the reverse effect on prices in unregulated markets.

A second consideration is the effect on risk. A price cap index that tracks short-term fluctuations in industry unit cost increases rate volatility but reduces utility operating risk. This can lower capital costs and/or permit an extension of the period between rate reviews that strengthens performance incentives.

A third consideration could be called implementation cost of the candidate index. This depends first on data availability. Data on price trends are available more quickly than the cost and quantity data that are needed, additionally, to measure TFP trends. Final data needed to compute the TFP growth of US power distributors in 2005, for instance, will not be available until late in 2006. The longer lag in the availability of cost and quantity data is due chiefly to the fact that these data typically come from *annual* reports whereas price indices are often calculated and reported on a *monthly or quarterly* basis. It is also germane that the calculation of TFP indexes can be quite a bit more complicated than the calculation of price indexes.

Implementation cost also depends on the feasibility of calculating current long run trends accurately. Methods have been developed to measure the recent long run trend in the TFP of the industry. For example, a sample period suitable for calculating the recent long run trend can be chosen using research on the drivers of TFP index volatility. The recent long run trend in an industry's TFP is, moreover, often if not always a good proxy for the *prospective* trend over the next several years. [57]

The use of historical data on industry input price trends to calculate the prospective future trend is more problematic. Industry input price indexes are often volatile. The calculation of an average annual growth rate thus depends greatly on the choice of the sample period. It can be difficult to reach consensus on what sample period would yield a long term input price trend. One reason is that research on the short run drivers of fluctuations in utility input prices is not well advanced. Absent a scientific basis for sample period selection, the choice of a sample period can engender controversy and raise the risk of PBR for utilities. Higher regulatory risk can raise the cost of funds and reduce thereby the net benefits of PBR.

Historical trends in input prices are, furthermore, sometimes poor predictors of the trends that will prevail in the near future. Suppose, by way of example, that there has been rapid input price inflation in the last ten years but that the expectation is for more normal inflation in the next five years. In this situation, regulators would presumably be loath to fix PCI growth at a rate that reflects the 10-year historical trend.

Examination of input prices in the power distribution industry suggests that they are somewhat volatile. Since power distribution is capital intensive, the summary input price index is quite sensitive to fluctuations in the price of capital. These result from fluctuations in plant construction costs and the rate of return on capital. Both of these components are more volatile than the general run of prices in our economy. For example, the rate of return on capital depends on the state of the economy and on expectations regarding future price inflation.[58] From the late 1970s through the mid 1980s, for instance, bond yields were far above historical norms due in large measure to inflation worries spurred by oil price shocks. They fell gradually for many years thereafter as concerns about inflation receded. More recently, long bond yields have been held down by efforts of the governments of China and other exporting countries to control exchange rates. Speculation on when and how much these policies will change is a staple of the financial press.

A sensible weighing of these three general considerations leads us to conclude that different treatments of input price and productivity growth are in most cases warranted when a PCI is calibrated to track the industry unit cost trend. The price inflation index should track short term input price fluctuations. The X factor, meanwhile, should generally reflect the long run historical trend of TFP.

This general approach to PCI design has important advantages. The inflation measure exploits the greater availability of inflation data. Making the PCI responsive to short term input

---

[57] Reliance on the long run trend can be problematic, however, when applied to utilities that contemplate major capital additions.
[58] The rate of return on capital also reflects return on equity. Returns on equity have also been volatile and are not highly correlated with bond yields.

price growth reduces utility operating risk without weakening performance incentives. Having X reflect the long run industry TFP trend, meanwhile, sidesteps the need for more timely cost data and avoids the chore of annual TFP calculations.

Given that the price inflation index should track recent input price growth, other important issues of its design must still be addressed. One is whether it should be *expressly* designed to track industry input price inflation as per relation [13]. We have noted several precedents for the use of an industry-specific inflation measure in rate adjustment indexes. However, the majority of rate indexing plans approved worldwide feature measures of economy-wide *output* price inflation such as the GDP-PI.

When a macroeconomic inflation measure is used, the PCI must be calibrated in a special way if it is to track the industry unit cost trend. Suppose, for example, that the inflation measure is the GDP-PI. This was noted above to be an index of output price inflation. Due to the broadly competitive structure of our economy, the long run trend in the GDP-PI is the difference between the trends in input price and TFP indexes for the economy.

$$trend\ GDPPI = trend\ Input\ Prices^{Economy} - trend\ TFP^{Economy}. \qquad [14]$$

Equations [12] and [14] together imply that

$$\begin{aligned} trend\ Unit\ Cost^{Industry} = {} & trend\ GDPPI \\ & - \left[\begin{array}{l} \left(trend\ TFP^{Industry} - trend\ TFP^{Economy}\right) \\ + \left(trend\ Input\ Prices^{Economy} - trend\ Input\ Prices^{Industry}\right) \end{array} \right] \end{aligned} \qquad [15]$$

When the GDP-PI is used as the inflation measure, it follows that the PCI already tracks the input price and TFP trends of the economy. X factor calibration is warranted only to the extent that there are differences in the input price and TFP trends of the utility industry and the economy.

This analysis suggests that when the GDP-PI is employed as a price inflation index the PCI can be calibrated to track the industry unit cost trend when the X factor has two calibration terms: a productivity differential and an input price differential. The productivity differential is the difference between the TFP trends of the industry and the economy. X will be larger, slowing PCI growth, to the extent that the industry TFP trend exceeds the economy-wide TFP trend that is embodied in the GDP-PI. The input price differential is the difference between the input price trends of the economy and the industry. X will be larger (smaller) to the extent that the input price trend of the economy is more (less) rapid than that of the industry.

The input price trends of a utility industry and the economy can differ for a number of reasons. One possibility is that prices in the utility industry grow at different rates than prices in the economy as a whole. For example, labor prices may grow more rapidly to the extent that utility workers have health care benefits that are better than the norm. Another possibility is that the prices of certain inputs grow at a different rate in some regions than they do on average throughout the economy. It is also possible that the industry has a different mix of inputs than the economy. Power distribution technology is, for example, noted above to be more capital intensive and therefore volatile in the price of capital than the typical production process in our economy.

The difficulties, discussed above, in establishing a long-term input price trend also complicate identification of an appropriate input price differential. For example, the difference between the average annual growth rates of input prices for the industry and the economy is sensitive to the choice of the sample period. It is less straightforward to establish the relevant sample period for a comparison of long-term industry and economy input price trends than it is for an analogous TFP trend comparison. Even if we could establish a differential between the long term trends it could differ considerably from the trend expected over the prospective plan period. This situation invites gaming over the sample period used to calculate the input price differential. Controversy is possible, additionally, over the method used to calculate the price of capital.

Logic of Revenue Cap Indexes

The extension of index logic to the case of revenue caps is straightforward. A revenue cap index that is based on index logic would be calibrated to track the cost trend of the industry rather than the unit cost trend. The cost trend of an industry is the sum of the unit cost trend and the output quantity trend. Recalling the results in [12] it follows that the cost trend is the difference between the input price and productivity trends plus the output quantity trend. An RCI is then calibrated to track the industry cost trend if

$$\begin{aligned} trend\ RCI &= trend\ Input\ Prices - trend\ TFP \\ &\quad + trend\ Output\ Quantities \end{aligned} \qquad [16]$$

The RCI growth formula thus differs chiefly from the PCI growth formula chiefly in considering a provision for output growth.

Precedents for Index Regulation

The earliest use of index logic in regulation design emerged from hearings before US federal regulatory commissions. As early as 1980, the Interstate Commerce Commission (ICC) proposed to determine allowable increases in rail freight rates using the average increase in rail carrier costs.[59] The Staggers Rail Act of 1980 was noted above to require index-based regulation for larger railroads. The law established a Zone of Rate freedom for certain rail services. Under section 203 of the Act, the boundary of this zone was to be adjusted each quarter by an "Index of Railroad Cost compiled or verified by the commission with appropriate adjustments to reflect the changing composition of railroad cost, including the quality and mix of material and labor". The growth rate of this index came to be called the Rail Cost Adjustment Factor (RCAF).

There was vigorous and protracted debate before the ICC regarding the appropriate form of this index. The most fundamental issue was whether the index should reflect the trend in the TFP of the industry as well as the input price trend. A TFP adjustment was opposed by railroads but favored by shippers. An index reflecting both would track the unit cost of the industry, as noted above. In 1989, the ICC concluded that the index should reflect the TFP trend of the railroad industry as well as its input price trend.[60] The X-factor it adopted was a moving average of the growth rate in an index of railroad industry TFP, as noted above.

Since the approval of the first plans at the federal level, rate-indexing plans have been adopted by a number of other regulatory commissions. The industry unit cost standard is frequently observed in PCI design. The US Federal Communications Commission has issued landmark decisions on PCI design that are broadly consistent with index logic. In approving the price cap plan for AT&T in 1989[61], inflation measures and industry TFP trends were discussed extensively.[62] The X-factor reflected the industry productivity trend and an inflation measure adjustment.

In approving rate indexing for the interstate services of local exchange carriers, the need to calibrate the PCI to the industry unit cost trend was explicitly recognized. For example, in a

[59] ICC, Advanced Notice of Proposed Rulemaking, "Railroad Cost Recovery Procedures," Ex Parte No. 290 (Sub-No. 2), April 28, 1980, 49 CFR 1135.
[60] ICC, "Decision, Railroad Cost Recovery Procedures-Productivity Adjustment," Ex Parte No. 290 (Sub-No. 4), March 22, 1989.
[61] "In the Matter of Policy and Rules concerning Rates for dominant Carriers." 4FCC Rcd 27763; CC Docket No. 87-313 (March 15, 1989).
[62] The affected rates of AT&T were subsequently decontrolled.

1995 order dealing with PCI, the FCC stated that "the indexes are adjusted each year in accordance with a formula that accounts for industry-wide changes in unit costs".[63] Another example was the Massachusetts Department of Public Utilities, in approving a rate-cap plan for the incumbent local exchange carrier, NYNEX, noted in 1995 that,

> Price cap regulation…replaces company specific test year cost based control of a firm's rates with an index representing the expected changes in the cost of the average firm in the industry.[64]

The California Public Utilities Commission noted in the same year in approving the rate-cap plan for Southern California Edison that

> The price and productivity values should come from national or industry measures and not from the utility itself. The independence of the update rule from the utility's own costs allows PBR regulation to resemble the unregulated market where the firm faces market prices which develop independently of its own cost and productivity. The productivity measure should come from a forecast of industry-specific productivity."[65]

In Canada, the CRTC has also subscribed to the industry unit cost standard. In an order approving the PBR plan for the Stentor companies, the CRTC stated that, "the price cap formula is composed of three basic components which, in total, reflect changes in the industry's long run unit costs.[66]

The price cap index approved by the OEB in the first generation plan for power distributors was constructed from industry specific input price and TFP trends. It was thus expressly designed to track the industry unit cost trend. Note, however, that the OEB did not elect to base the X-factor solely on the *long-term* TFP trend.

As for RCI logic precedents, the RCI approved by the OEB for the O&M expenses of Enbridge Gas Distribution was based on index logic. So too was the revenue per customer index approved by the CPUC for Southern California Gas.

---

[63] Federal communications Commission, First Report and Order in the Matter of Price Cap Performance for Local Exchange Carriers, cc Docket 94-1, April 7, 1995.

[64] Petition of New England Telephone and Telegraph Company dba/NYNEK for an Alternative Regulatory Plan for the company's Massachusetts Intrastate Telecommunications Services. DPU 94-50. May 12, 1995.

[65] Application of Southern California Edison to adopt a Performance Based Rate Making Mechanism Effective January 1, 1995, Alternate Order of Commissioners Fessler and Duque, July 21, 1996.

[66] Ibid paragraph 29.

Productivity Measurement

What is Productivity? Since productivity plays an important role in North American rate indexing some discussion is warranted as to how it is calculated and what is known about productivity trends. Productivity is conventionally measured using a productivity index. The productivity index of an industry is the ratio of indexes measuring its output and input quantities.

$$TFP = \frac{Output\ Quantities}{Input\ Quantities}. \qquad [17]$$

An output quantity trend index for an industry summarizes trends in the amount of work that it performs. If output is multidimensional, the growth in each output quantity dimension considered is measured by a subindex. The growth in the index is a weighted average of the growth in the quantity subindexes.

The productivity index of an industry captures the wide range of developments that can cause its unit cost to grow at a different rate than its input prices. Productivity is volatile from year to year but typically trends upward over periods of several years. The unit cost trend of an industry will then typically be slower than its input price trend.

An input quantity trend index for an industry summarizes trends in the amounts of production inputs that it uses. Growth in the usage of each input category considered is measured by a subindex, e.g. a labor quantity subindex. The growth in the summary input quantity index is typically a weighted average of the growth in the quantity subindexes.

Productivity indexes are often classified by the scope of the inputs that they cover. Thus, a labor productivity index addresses only productivity in the use of labor inputs. A multifactor productivity index measures productivity in the use of multiple categories of inputs. A TFP index measures productivity in the use of all inputs that a company uses.

Sources of Productivity Growth Mathematical and econometric research has been conducted for two decades on the sources of productivity growth. One important source is change in production technology. Technological change slows cost growth and permits an industry to produce given output quantities with fewer inputs.

Economies of scale are a second source of productivity growth. These economies are available in the longer run if cost characteristically grows less rapidly than output. In that event, output growth can slow unit cost growth and raise productivity. A company's potential for scale economy realization depends on its current operating scale and on the pace of its output growth.

Incremental scale economies will be greater the more rapid is output growth and the smaller is the initial operating scale. Utilities of sufficient scale may no longer realize scale economies from output growth and may even experience diseconomies that slow productivity growth.

A third determinant of productivity growth is changes in the miscellaneous business conditions other than input prices and output quantities that affect cost. Generally speaking, a business condition trend that tends to lower cost will accelerate productivity growth. A business condition that raises cost will tend to slow productivity growth.

Pacific Economics Group has done extensive econometric research over the years on the determinants of power distribution cost. An article on our research was recently published in the *Energy Journal,* the refereed journal of the International Association of Energy Economists.[67] We have found that the total cost of power distribution tends to be higher:

- The greater is the extent of system undergrounding
- The greater is the number of gas distribution customers served
- The greater is system age
- The greater is the percentage of deliveries made to residential and commercial customers
- The greater is the ruralization of the service territory served
- The greater is the forestation of the service territory
- The greater is the diversification into generation and transmission

Changes in these conditions can affect cost and productivity growth. For example, growth in the number of gas customers served by a utility can slow growth in its power distribution cost and accelerate productivity growth.

A fourth determinant of productivity growth is X inefficiency. This is the degree to which individual companies operate at the maximum efficiency that technology allows. Productivity will grow (decline) to the extent that X inefficiency diminishes (increases). The potential of a company for productivity growth from this source is greater the greater is its current level of operating inefficiency.

An important source of productivity growth in the shorter run is the degree of capacity utilization. Producers in most industries find it uneconomical to adjust production capacity to

[67] Mark Newton Lowry, Lullit Getachew, and Dave Hovde, “Econometric Benchmarking of Power distribution Cost, *Energy Journal*, July 2005

short-run demand fluctuations. The capacity utilization rates of industries therefore fluctuate. productivity grows (declines) when capacity utilization rises (falls) because output is apt to change much more rapidly than capacity.

Another short-run determinant of productivity growth is the pattern of expenditures that are occasional in character. Expenditures of this kind include those for certain kinds of maintenance and investments. A surge in expenditures can slow productivity growth and even result in a productivity decline. Uneven spending is one of the reasons why the productivity growth of individual utilities is often more volatile than the productivity growth of the corresponding industry.

Econometric Measurement of Productivity Growth The mathematical analysis and econometric research that we use to investigate the determinants of productivity growth also make possible an alternative econometric approach to the measurement of productivity trends. A notable disadvantage of this approach is its complexity. A notable advantage is the ability to customize productivity trends to the special operating conditions of individual utilities without losing the external character of the research that is so valuable in the design of rate escalation indexes. We can, for example, calculate productivity trends for individual Ontario utilities that are specific to their operating scale and their expectations concerning output growth, undergrounding, and other business conditions. This general approach has occasionally been used in California regulation. For example, CPUC staff has made estimates of the expected productivity growth of individual utilities that are specific to their operating scale.

Productivity Precedents The productivity trend of a utility industry is an empirical issue. Results of productivity research have been presented in several PBR proceedings. Regulators often choose X-factors without stating their views on the components. There are, however, several cases in which they have explicitly acknowledged the long run industry productivity trend. Table 1 presents a summary of the North American precedents, along with some additional information about X factor precedents that is discussed further below. Table 2 reviews some precedents from the telecommunications industry.

Table 1
X FACTORS APPROVED BY NORTH AMERICAN REGULATORS FOR GAS AND ELECTRIC UTILITIES

Industry	Company	Term	Jurisdiction	Acknowledged Productivity Trend	Inflation Measure	Stretch Factor	X-Factor	Comments
Gas distribution	Boston Gas (I)	1997-2003	Massachusetts	0.40%	GDPPI	0.50%	0.50%	
Gas distribution	Boston Gas (II)	2004- 2013	Massachusetts	0.58%	GDPPI	0.30%	0.41%	
Gas distribution	Berkshire Gas	2002-2011	Massachusetts	0.40%	GDPPI	1.0%	1.0%	Adopted the productivity study used by Boston Gas I
Gas distribution	Consumers Gas	2000-2002	Ontario	0.63%	CPI	0.50%	1.10%	O&M Productivity
Gas distribution	Union Gas	2001-2003	Ontario	0.9%	GDPPI	0.5%	2.5%	
Gas distribution	San Diego Gas and Electric	1999-2002	California	0.68%	Industry specific	0.55% (Average)	1.23% (Average)	
Gas distribution	Southern California Gas	1997-2002	California	0.50%	Industry specific	0.80% (Average)	2.30% (Average)	Special 1% factor added to X to reflect declining rate base
Gas distribution	Bay State Gas	2006-2015	Massachusetts	0.58%	GDPPI	0.4%	0.51%	Adopted Boston Gas II
Bundled power service	Pacificorp	1994-1996	California	1.4%	Industry specific	NA	1.4%	Company specific productivity
Power distribution	San Diego Gas and Electric	1999-2002	California	0.92%	Industry specific	0.55% (Average)	1.47% (Average)	
Power distribution	Southern California Edison	1997-2002	California	NA	CPI	0.58% (Average)	1.48% (Average)	0.90% productivity trend estimated by Edison and Commission staff but not formally acknowledged by CPUC
Power distribution	All Ontario distributors	2000-2003	Ontario	0.86%	Industry specific	0.25%	1.5%	Productivity trend referenced is the 10 year average growth rate X factor is based on 5 and 10 year weighted average
Power distribution	Nstar	2006-2012	Massachusetts	NA	GDPPI	NA	0.63% (average)	
Bundled power service	Central Maine Power (I)	1995-1999	Maine	NA	GDPPI	NA	0.9% (average)	
Power distribution	Central Maine Power (II)	2001-2007	Maine	NA	GDPPI	NA	2.57% (average)	
All utilities	**Sample Average**			**0.70%**			**1.21%**	
All industry specific	**Sample Average**						**1.58%**	
All macroeconomic	**Sample Average**						**1.01%**	

Table 2

Productivity Decisions in Telecommunications Regulation

Industry	Company	TFP Trend
Telecommunications	Canadian telcos	2.6
Telecommunications	SNET – CT	2.1
Telecommunications	Ameritech – IL	1.3
Telecommunications	Nynex – ME	2.2
Telecommunications	Nynex – MA	2.0
Telecommunications	Ameritech – OH	2.8
Telecommunications	Bell Atlantic – PA	2.9

Here are some salient findings from Table 1.

- 15 approved plans for energy utilities were examined which featured rate escalation indexes based, in whole or in part, on input price and productivity research. The jurisdictions involved were California, Massachusetts, Maine, and Ontario.
- While productivity evidence was part of the foundation for all of these X factors, there were several instances in which no explicit findings were made by regulators concerning productivity trends.
- The average of the acknowledged long-run TFP trends for gas and electric utilities were, respectively, 0.58% and 1.06%, respectively.

A comparison of the results in Tables 1 and 2 suggests that X-factors can reasonably be expected to be much higher in indexing plans for telecom services than in plans for many energy services. The recent TFP trend for telecom utilities has apparently been around two hundred basis points higher than that for energy distributors. This reflects, in the main, burgeoning technological change and demand growth in the telecommunications industry. It should not be surprising, then, to find approved telecommunications price cap plans with X-factors at least two hundred basis points above those in approved energy utility plans.

These productivity figures also help to explain why a multi-year rate freeze may not financially stress telecom utilities as much as it would an energy utility. Given input price growth in the 2-3% range, index logic suggests that telecom utilities have recently experienced steady or moderately declining unit costs. This permits them to prosper under rate freezes.

On the other hand, while energy utilities face an input price growth trend broadly similar to that of telecom utilities, their TFP growth is typically much slower. Accordingly, their input price growth is more likely to exceed their TFP growth, and their unit cost is more likely to rise over time. This is a common situation in our economy as can be seen by the tendency of consumer price indexes to rise over time. Many energy utilities will therefore have difficulty remaining financially viable for an extended period of time without nominal rate increases. An American-style PCI could address this situation by allowing utility rates to rise moderately each year in nominal terms to keep pace with industry unit cost growth. The fact that utility prices are apt to rise in nominal terms should by itself cause no more concern than in competitive sectors of the economy.

Regional Research Focus An important issue in North American style index calibration is the choice of a region for indexing research. Regions in different countries can exhibit different input price and productivity trends even if they are adjacent. Different regions within countries of some size can also exhibit different input price and productivity trends. There are evident differences in regional economic growth in both Canada and the US. Differences in government policies can lead to differences in the unit cost growth of utilities. For example, governments can differ in support for retail competition and/or demand-side management efforts that affect volume growth.

This analysis suggests that the region surrounding a utility will tend to have more similar input price and productivity trends than regions further afield. These considerations suggest that the unit cost trend in the region surrounding the subject utility can be the appropriate focus of input price and productivity research. However, circumstances can render this option unworkable as well. Some or all of the surrounding region may be in a different country. Additionally, the surrounding region may have few peer utilities, lack good utility operating data, or be dominated by just one or two utilities.

North American indexing plans recognize the region selection challenge. In Ontario, regulators elected to base the inflation and X-factors in the rate cap plan for power distributors on the input price and productivity trends of the provincial industry. The Massachusetts regulator explicitly approved the calibration of the Boston Gas X-factor using the TFP trend of northeast distributors. In California, one utility has used company-specific productivity evidence to support PBR plan design, whereas others have used evidence of national productivity trends.

In the telecommunications industry, X-factors in a number of telecommunications price cap plans for US LECs have been established in proceedings where the company's own productivity trend was the featured evidence. The FCC based its X-factor for interstate services of LECs on national TFP research but has acknowledged its potential inappropriateness for certain regions. The CRTC based its X-factor for LECs on national data in its first price cap plan.

Stretch Factor

A stretch factor is sometimes added to the X-factor of a North American-style index to improve the terms of service for customers. This will typically be larger to the extent that inefficiencies in the companies operations are known to be large and have not been fully dealt with in setting initial rates. In the absence of such evidence, extensive inefficiency cannot be assumed and stretch factors are typically set at some modest standard level.

An early use of stretch factors was in the initial price cap plan approved by the FCC for AT&T. A stretch factor of 0.5% was added to the calculated TFP differential of 2.5% to yield an X-factor of 3%. Since then, stretch factors have been featured in many North American indexing plans. They are sometimes explicit and sometimes implicitly added to the X-factor.

Table 1 summarizes cases in which North American commissions have approved explicit stretch factors. Table 3 summarizes precedents for telecom utilities

Table 3

STRETCH FACTOR PRECEDENTS FOR TELECOMMUNCATIONS UTILITIES

Industry/Company	Period	Stretch Factor
Ameritech - IL	1995-2002	1.0%
Ameritech - IL	2003-	1.0%
Ameritech - OH	1995-2003	0.2%
Bell Atlantic - PA	1995-2003	0.0%
Canadian Telecoms	1998-2001	1.0%
Canadian Telecoms	2002-	0.0%
NYNEX - MA	1995-2001	1.0%
NYNEX - ME	1995-2001	1.0%

It can be seen that

- The average value of the explicitly approved stretch factors for energy utilities was 0.54%.
- Stretch factors for telecom utilities are typically above those for energy utilities. This may reflect in part the more rapid TFP growth that typifies the telecom industry.

Note also that there is one case in which a commission has chosen an explicit stretch factor for a PBR plan at the conclusion of a previous plan. In that case, the CRTC elected to eliminate the stretch factor, stating that

> The commission agrees with the Companies' view that current price already reflect the impact of the stretch factor established in the initial regime. In additional, the Commission considers that additional productivity gains due to the further streamlining of regulation would be difficult to achieve in the next price cap regime. The Commission is also of the view that the basic productivity offset of 3.5%, based on the marginal cost approach, indirectly incorporate a limited stretch factor. This implicit stretch factor results from the fact that the marginal cost growth for the years 1998 to 2001 included the productivity achieved under price cap regulation. Accordingly, the Commission concludes that no stretch factor should be applied to the productivity offset.[68]

<u>Other X Factor Findings</u>

Additional information is available from several sources that is germane to the selection of X factors for energy distribution utilities.

X Factor Precedents Table 1 provides some additional information about X factor precedents that merits note.

The average approved X factor, gross of all adjustments, for the rate escalation indexes listed in Table 1 was 1.30%.

[68] CRTC, "Regulatory Framework for Second Price Cap Period", CRTC 2002-34, May 2002.

- The average approved X factor, gross of all adjustments, for rate escalation indexes featuring an industry-specific inflation measure was 1.58%. The relatively high X factor for SoCalGas is clearly an outlier. It reflects, in large measure, an unusual 100 basis point adjustment that the Commission made for an anticipated decline in the company's rate base.
- The average X factor, gross of all adjustments, for the rate escalation indexes featuring a macroeconomic inflation measure such as the GDP-PI was 1.01%.

Gas Distribution Rate Index The US Bureau of Labor Statistics computes producer price indexes (PPIs) for a wide range of goods and services that are offered to business establishments. Since 1991, a PPI has been computed for the unbundled transportation services that natural gas distributors offer to these establishments. The trend in this index is detailed in Table 4. It can be seen that over the full 1991-2005 sample period for which data are available, distribution rates rose at a 2% average annual rate. The full sample period encompasses sub-periods in which the rate trends were quite divergent. From 1995 to 2000, for instance, rates actually fell by an average of 0.6% annually. Over the 2000-2005 period, in contrast, rates rose at a 2.3% annual pace.

Table 4 also computes the X factors that would produce such rate escalation in a PCI with a GDP-PI – X growth rate formula. It can be seen that over the full sample period, the implicit X factor was 0.7%. Over the most recent ten years, the implicit X factor was 1.1%. Over the most recent five years, the implicit X factor was 0.0%.

Table 4

IMPLICIT X FACTOR IN GAS DISTRIBUTION RATES, 1991-2005

	PPI Natural Gas Distribution - Transportation Only		GDP-PI		
Year	**Level**	**Growth Rate**	**Level**	**Growth Rate**	**Implied X Factor**
1991	96.8		84.5		
1992	99.5	2.8%	86.4	2.3%	
1993	101.5	2.0%	88.4	2.3%	
1994	101.2	-0.3%	90.3	2.1%	
1995	106.9	5.5%	92.1	2.0%	
1996	105.7	-1.1%	93.9	1.9%	
1997	109.4	3.4%	95.4	1.6%	
1998	103.6	-5.4%	96.5	1.1%	
1999	102.3	-1.3%	97.9	1.4%	
2000	103.9	1.6%	100.0	2.2%	
2001	103.4	-0.5%	102.4	2.4%	
2002	105.5	2.0%	104.2	1.7%	
2003	108.2	2.5%	106.3	2.0%	
2004	113.3	4.6%	109.1	2.6%	
2005	116.3	2.6%	112.2	2.8%	
Formula		**[B]**		**[A]**	**[A] - [B]**
Average 91-05		**1.3%**		**2.0%**	**0.7%**
Average 95-05		**0.8%**		**2.0%**	**1.1%**
Average 00-05		**2.3%**		**2.3%**	**0.0%**
Average 95-00		**-0.6%**		**1.6%**	**2.2%**

Source, PPI Natural Gas Distribution Transportation Only: Bureau of Labor Statistics; http://www.bls.gov

Source, GDP-PI: Bureau of Economic Analysis; http://www.bea.gov

Note: Assumes GDPPI - X Index Formula

Rate Freezes Rate freezes have, as noted above, been approved over the years which apply to the distribution services of a number of companies. These precedents should be interpreted cautiously in the selection of X factors, for several reasons.

- A number of freezes on electric utility rates have been occasioned by power market restructuring agreements. These freezes may reflect company-specific considerations of appropriate stranded generation cost recovery and the amount of rate stability needed in the run-up to competition. When a freeze pertains to a bundle of services that includes power supply it is, furthermore, apt to be more sensitive to the expected trend in the unit cost of generation than to the trend in the unit cost of distribution. The trend in the unit costs of generation and distribution can differ substantially over the five to seven year periods that have been typical of restructuring-related freezes. For example, investments in generation are distributed less evenly over time than those for distribution and the unit cost of generation can grow quite slowly after major plant additions are no longer anticipated. Some freezes apply to all rate components while others permit adjustments to distribution rates under the cap. In California, where adjustments to distribution rates are allowed, they have escalated considerably despite the overall cap.
- Many other rate freezes are components of merger agreements. Examples include recent freezes pertaining to the power distribution services of NSTAR and National Grid in Massachusetts. Rate freezes are made possible under good mergers by a temporary but unsustainable surge in productivity growth.

4.2.2 The British Approach to Index Design

The British approach to the design of rate escalation indexes is so-called because it is typical of utility regulation in Britain. It has since been adopted in several other countries. Most notable, perhaps, is its widespread use in Australia.[69] Most British utilities were formerly public

[69] Other countries that have used the British approach to inexing include Ireland and Mexico.

enterprises. British Telecom (BT) was the first to be privatized, in 1984. Since then, privatization has extended to Britain's electric, gas, and water utilities.

The decision to use rate indexing in British utility regulation was strongly influenced by the recommendations of Stephen Littlechild of the University of Birmingham. In a report released in 1983, he proposed to adjust BT's rates using an index with a growth rate formula of "RPI-X" form.[70] A specific value for X was not recommended, nor was there significant discussion in Littlechild's paper of the appropriate framework to be used to determine X. Rather, the value for X was described as "a number to be negotiated." The lack of a well-defined framework has given British regulators considerable discretion in determining X-factors. Over time, however, broadly similar approaches have developed across Britain's utility industries.

Under "British-style" indexing rate cases are typically held every five years. In contrast to North American practice, which focuses on a single test year, the rate case involves detailed multi-year cost and output forecasts. The principle "building blocks" of the total cost forecast are the forecasts of the value of the current capital stock and of capital spending, depreciation, the rate of return on capital, and O&M spending. A macroeconomic inflation index such as the RPI is used as the inflation measure of the price cap index. Given the forecasts of growth in total cost, volumes and other billing determinants, and the RPI, it is possible to choose a combination of initial rates and an X-factor that equates forecasted revenue and forecasted cost.

The British approach to the design of rate and revenue cap indexes has several advantages over the North American approach. One is the ability to implement it in situations where the North American approach is hampered by a lack of historical data that could be used for productivity calculations. This was, apparently, the situation in most of the British industries at the time of their privatization. The British approach is also advantageous in a situation where there really is no sizeable group of peers that could provide the basis for industry productivity trends even if data were available. This continues to be the situation in the British power and natural gas transmission industries.

The British approach is also advantageous in situations where the expected forward looking productivity trends of individual utilities are markedly different from the recent long-run TFP trend of the industry. This situation is often encountered in industries, like power

[70] Stephen Littlechild, "Regulation of British Telecommunications' Profitability: Report to the Secretary of State",mimeo, February 1983.

generation and transmission, that have unusually bunched intertemporal patterns of investment. In that case, an individual utility might, for example, anticipate large scale investments in the next few years that will slow productivity growth markedly even though the recent productivity growth of the industry is fairly rapid.

These advantages of British-style indexing should be weighted against some important disadvantages. One disadvantage is the higher level of regulatory cost that it involves. A five year test rate case is substantially more complicated than the single test year cases that go hand in hand with productivity indexing. The uncertainties of long term forecasts may also be said to discourage plans with unusually long terms. Another serious problem with the British approach is the incentive that it provides to utilities to exaggerate their future cost growth.

These disadvantages have spurred considerable innovation in British style regulation in recent years.

- Statistical benchmarking is frequently used to appraise O&M expense forecasts.
- A consultant is hired by OFGEM to provide an independent assessment of the company's capex needs.
- A predetermined formula establishes an allowed level of capex that is a function of the utility's and the consultant's capex recommendations. If the utility's recommendation is below or even close to the consultant's, the budget is set *above* the consultant's recommendation as a reward. When the utility's budget is higher than the consultant's, on the other hand, the allowed capex is set between the two proposals. The greater is the disparity between the proposals, the closer is the allowance set to the consultant's proposal.
- There is, additionally, a formula for sharing any deviation of actual capex from the capex budget. The company share is greater (to its benefit) the lower is its capex proposal.
- The Essential Services Commission in Victoria, now uses industry productivity research to forecast future O&M expenses in a rate setting system that is otherwise British in character.

5 SERVICE QUALITY PROVISIONS

The attainment of appropriate quality standards is a critically important consideration in PBR plan design. Utilities can often save money by trimming maintenance expenditures and capital investments that affect quality. The threat of lost business is weaker for utilities than for other businesses where product quality is a vehicle for competition. In many cases, the local utility is a monopoly provider and stands to lose fewer sales than a competitive firm if service quality is off the mark. Regulation may also deny the utility the flexibility it needs to offer different price-quality mixes.

The OEB noted the importance of service quality oversight in its decision on the first *Rates Handbook*. It stated that

> Any reduction in the quality and/or reliability of service represents a reduction in the value of that service. Therefore, as part of its function in regard to approving or fixing just and reasonable rates, the Board has a responsibility to oversee that service quality is preserved and improved.[71]

Formal service quality incentive mechanisms have been approved for numerous utilities. They are a form of benchmark PBR which rewards or penalizes a utility depending on the relationship between its measured quality of service and quality benchmarks. There are three basic elements in a service quality incentive plan: a series of indicators of the company's quality of service; an associated set of quality benchmarks; and an award mechanism that leads to changes in utility rates or allowed returns. The indicators are measurable service quality dimensions. The benchmarks are the standards against which the indicators are judged. They can in principle be based on the company's historical performance, industry norms, or levels that are deemed to be acceptable for other reasons. The award mechanism determines the adjustment in rates that is warranted by the change in service quality. Important design issues include the symmetry of awards and penalties and the customers' valuation of specific quality indicators.

5.1 Benchmarking Basics

Benchmarking mechanisms involve the evaluation of one or more indicators of company activity using external performance standards (benchmarks). The standards are external to the extent that they are insensitive to the actions of subject utility managers. Evaluations and rate adjustments are accomplished by formal mechanisms that are established in advance of use and typically function for several years.

The key features of a benchmark plan are the performance indicators, performance benchmarks, and the rate adjustment mechanism. The performance indicators used in approved benchmark plans vary greatly in scope. Plans are comprehensive to the extent that they cover all of the utility performance dimensions that matter to customers.

The performance benchmarks used in benchmark plans are also varied. A common benchmark is a company's activity level in a period just prior to plan commencement. A company is then rewarded for improvement in its performance relative to recent history.

An alternative approach, which is an example of "yardstick regulation" or statistical benchmarking, is to use the corresponding performance indicator of a group of utilities. Under this approach, a company is rewarded for improving its performance indicator relative to the group. The utility group is sometimes called a peer group, but can consist of all utilities in the same region as the company subject to the plan. In that event, the peer group may be viewed as a proxy for the regional industry. In principle, the region can also be the entire nation.

The rate adjustment mechanisms in approved benchmark plans vary. A major design issue is the customer sharing percentage. The mechanism may or may not feature a deadband in which deviations from the benchmark do not induce rate adjustments.

Benchmarking plans provide supplemental adjustments to rates rather than serving as the sole basis for rate adjustment restrictions. Several rate adjustment mechanisms can, in principle, coincide with a benchmarking plan. At one extreme, rates may be adjusted for the actual trend in a company's unit cost. At the other, rates may be predetermined (set in advance) for several years.

5.2 Quality Indicators

A critical issue in the development of an effective service quality provisions is the choice of indicators on which performance will be judged. Ideally, individual quality indicators should satisfy four criteria: 1) They should be related to the relevant aspects of service; 2) focus on monopoly services; 3) cover all major quality dimensions; and 4) be no more complex than necessary to provide effective incentives.

---

[71] OEB, *Rates Handbook Decision*, ibid p. 50.

First, since measured service quality can ultimately affect customer rates, indicators should be linked to aspects of utility service customer's value. This may seem obvious, but a strict application of these criteria excludes indicators that have been included in some plans. For instance, the knowledge and courtesy of phone center employees may be a legitimate quality indicator, but the goal of establishing worker training programs to build these skills is not.

Second, indicators should focus on the quality of the activities for which there are few if any alternative suppliers. This is consistent with the principle that regulation, including regulation of service quality, is less necessary in competitive markets. Market forces are likely to create acceptable quality levels when products are available from multiple providers. Third, quality indicators should not focus on some important areas while ignoring others because performance may deteriorate in the non-targeted areas.

Comprehensiveness can be achieved simply by adding indicators to a plan. However, regulatory costs often rise accordingly since more utility and commission resources must be devoted to quality monitoring, measurement, and the reconciliation of findings related to quality indicators. Some commissions have been sensitized to the regulatory costs of complex service quality plans. In these jurisdictions, service quality incentives have been simplified by relying on fewer, but more broadly-based, indicators. While the specific indicators may vary widely among approved service quality incentive plans, there are broad similarities between the types of indicators used for energy utilities.

5.3 Quality Benchmarks

Quality benchmarks are the standards against which measured quality is judged. Benchmarks should be ideally sensitive to a utility's external business conditions, which influence quality and are relatively immune to the influence of random events. These business conditions may be called quality "drivers." The list of relevant factors includes weather (*e.g.* winds, lightning, extreme head and cold), vegetation (contact with power lines), the amount of undergrounding mandated by local authorities, the degree of ruralization in the territory (typically increasing the exposure of feeders to the elements and lengthening response times when faults occur), the difficulty of the terrain served and regulatory changes such as a restructuring of the industry to promote competition. These drivers can vary considerably between utilities and over time.

Universally accepted quality standards do not exist for utility industries, so commissions have considerable latitude in setting benchmarks. For any given indicator, one straightforward benchmark is the utility's average performance over a recent period. Quality assessments would then depend on measured quality levels that differ either positively or negatively from recent historical experience.

Using past utility performance to set benchmarks is appealing in many ways. The data are of known quality and reflect local cost drivers. The construction of benchmarks from a utility's past quality level should reflect the fact that a company's measured service quality performance can be affected by external business conditions that are beyond management control. Some of these business conditions are volatile and prone to fluctuations that are hard to predict. Utilities should not ideally be subject to penalties or rewards because random factors have affected their measured service quality. PBR plans can be designed to mitigate the impact of random factors in leading to inappropriate penalties or rewards.

One way to handle the impact of fluctuations in quality drivers is through a deadband around the quality benchmark in the award/penalty mechanism. Statistical methods can provide a rigorous foundation for setting deadbands that reduce the probability of inappropriate penalties or rewards to specified levels (*e.g.* 5%). Such statistical methods have been used in several service quality PBR plans for telecom utilities and have been proposed by energy utilities in some states.[72]

Statistically based dead bands should reflect historical fluctuations in indicator values. This is commonly measured by the standard deviation of sampled values. The greater the fluctuations have been, the higher the standard deviation and the wider the deadbands. Statistically based deadbands may in principle reflect the size of the sample. The deadband should be wider the smaller is the sample.

Regulators may not consider a utility's past performance to be an adequate quality standard, especially if recent service levels were deemed poor. Some utility managers may also view the company's history as inappropriate when its performance is exceptionally good. In this case, it may be considered unfairly demanding to expect the utility to match its historically superior performance on an ongoing basis.

An alternative to basic benchmarks on the Company's own history is to base them on the service quality performance of the industry. The industry may take the form of a national or regional sample or a peer group selected by other means. In principle, industry-based benchmarks may be attractive in PBR. They are clearly external to the subject utility, which creates strong performance incentives. Industry benchmarks also tend to be consistent with the operation of competitive market, where customer choices are driven by the cost and quality of products relative to available substitutes.

In practice, however, industry-based benchmarks are often problematic. One reason is that uniform and publicly-available data are not collected for large numbers of energy utilities. This lack of available data probably explains why so few approved plans contain industry-based quality benchmarks. While this is a recognized problem, some commissions (*e.g.* Massachusetts) are nevertheless examining the desirability of using peer data within their state to set reliability benchmarks for individual utilities.[73]

Another reason that industry-based benchmarks are problematic is differences in the operating conditions of utilities. Optimal quality levels reflect such key conditions as the cost of providing quality service and the demand for quality. These conditions vary across service territory. The issue of key importance is whether a company's quality level is good given the quality drivers that it faces. It is difficult to obtain a sizable amount of quality data from companies that are similarly situated.

5.4 Award and Penalty Rates

Another significant plan design issue is the magnitude of any rewards or penalties levied. In practice, empirical evidence is rarely presented to justify the amount of potential penalties or rewards in a plan. Instead, penalty levels are sometimes chosen with the idea that they are "significant" enough to prevent service quality declines. The rationale seems to be that the penalties should at least exceed cost savings that the utility might expect by cutting resources used to deliver service quality.

---

[72] "Investigation by the Department of Telecommunications and Energy on its own motion to establish guidelines for service quality standards for electric distribution companies and local gas distribution companies." Massachusetts D.T.E. 99-84 (June 29, 2001).

[73] "Investigation by the Department of Telecommunications and Energy on its own motion to establish guidelines for service quality standards for electric distribution companies and local gas distribution companies." Massachusetts D.T.E. 99-84 (June 29, 2001).

Ideally, a service quality incentive requires information on how customers value different quality indicators, so that the potential rewards and penalties for performance will reflect the value of the service provided. Given its importance, it is somewhat surprising that little empirical work has been done on customer valuations of quality indicators included in incentive plans. In part this is because quality is inherently difficult to value. But while this information may not be readily available, it can be gathered from a number of sources.

Although a complete discussion of the topic is beyond the scope of this article, three basic methods are used to estimate the value of service quality. One method uses proxy data related to the service attribute. For example, the value of having to wait for a field service representative to arrive can be approximated as the customer's lost wages (*i.e.*, the opportunity cost of the customer's time). Proxy prices have the advantage of simplicity, but they can be imprecise and bear a tenuous link to actual service valuations.

A second method of estimating customer valuation uses market-based measures for the value of service. The difference between firm and interruptible rates is one example of market-based data that reflects some customers' valuations of reliability. Another example of market-based measures is the use of hedonic price indexes, which are developed by regressing market prices on identifiable quality attributes. Hedonic price indexes reflect the notion that price differences are due to implicit markets for individual product characteristics. Some official statistics utilize hedonic methods; for example, the Bureau of Labor Statistics adjusts for quality changes of some products when computing the Consumer Price Index. While market-based methods are often conceptually sound, they can be controversial, are often not well-understood, and can produce divergent estimates of underlying quality valuations. In addition, hedonic methods are less likely to capture the underlying quality valuations in utility markets since prices often reflect regulatory decisions rather than market forces.

Finally, quality valuations can also be obtained through customer surveys. An advantage of this approach is that surveys can focus on specific aspects of utility services that might be included in an incentive plan. However, survey results reflect subjective perceptions rather than actual consumer behavior, and hypothetical valuations may not be a good guide to how consumers would actually act in markets.

5.5 Plan Symmetry

The symmetry of the award mechanism is another important design issue. It has been argued that symmetric awards (*i.e*. both rewards and penalties are possible) are not needed when quality incentives are designed only to maintain quality levels which might otherwise decline due to the stronger incentives to cut costs under PBR. However, symmetric plans can be calibrated to incent only the maintenance of current quality standards.

The encouragement of better quality may, in any event, be desirable. All types of PBR, including service quality incentives, are fundamentally motivated by a desire to improve utility performance and not simply to prevent performance from slipping. Asymmetric plans generally do not create incentives for companies to improve quality and thus may limit the total customer benefit that is available from utility operations.

The impact of external business conditions on measured service quality performance also tends to support symmetric service quality incentives. As noted, some business conditions can be quite volatile and may lead to inappropriate penalties or rewards. Symmetric service quality incentives reduce the likelihood that random factors will lead to inappropriate net penalties or rewards over the course of a multi-year incentive plan. That is because random changes in business conditions can lead to rewards as well as penalties. Over time, the magnitudes of any inappropriate penalties and rewards can therefore be expected to cancel each other out. This leads to reasonable penalties and rewards that on average reflect a utility's underlying quality performance. This would not be the case with an asymmetric service quality incentive, where external factors may subject a company to penalties without the chance of being compensated with offsetting rewards.

Symmetric plans are also more consistent with the workings of unregulated markets. Customers in such markets routinely pay higher prices for higher quality products. Many farmers, for instance, do not have full control over the quality of their produce from year to year and earn quality premia when production conditions are favorable as well as lower prices when they are unfavorable.

However, competitive markets usually offer an array of goods with varying quality levels, and not all customers choose to consume high-quality goods. In some cases, incentive plans lead to price increases on monopoly services. Where this is the case, at least some customers may be paying for quality improvements that they do not want.

The uncertainties related to the magnitude of rewards or penalties lend additional support for symmetric service incentives over asymmetric incentives. Since regulators often use considerable discretion in setting penalty rates, a symmetric plan may discipline regulators into choosing more appropriate rates. That is, with an asymmetric plan, regulators may err on the side of choosing very high penalties to assure that quality does not decline under the plan. This is less likely under a symmetric plan, which would require an equally high reward due to performance improvements. Hence, even if an asymmetric plan is ultimately approved, a symmetric service quality proposal may be beneficial if the prospect of symmetry leads to more appropriate magnitudes for penalty payments.

5.6 Precedents

There are a large number of formal service quality provisions in approved rate plans. Service quality PBR is especially well established in New York and California. Generic proceedings on service quality PBR have been held in several states.[74] A proceeding of this kind is now underway in Ontario.

Symmetric service quality plans have been approved for energy utilities. For example, both the California and New York commissions have adopted symmetric service quality plans based on explicit findings that the underlying principles are sound. However, asymmetric service incentives are somewhat more common.

Despite the many precedents for formal service quality incentive mechanisms, many PBR plans do not have them. The absence of incentive mechanisms is especially common in first generation plans. For example, the OEB did not approve a formal mechanism for power distribution in its Rates Handbook decision. It stated in the decision that

> *The Board recognizes that electricity industry restructuring introduces many unknown factors that could impact on performance levels and customer expectations. Further, there is a lack of consistent information on historical performance. Therefore, the Board is of the view that, for first generation PBR, a cautions approach to introducing service quality performance indicators and standards is warranted. The proposed approach in first generation PBR appropriately focuses on data collection, reporting, and monitoring of service quality and reliability performance by all distribution utilities.*[75]

---

[74] See, for example, Massachusetts D.T.E. 99-84, op cit.

[75] Ontario Energy Board. *Decision with Reasons*, RP-1999-0034. Jaunary 18 2000, p. 50.

The Board also elected not to approve a formal quality incentive mechanism in the first general Union Gas PBR plan.

5.7 Informal Quality Provisions

Service quality PBR is becoming more important in utility regulation. Quality incentive mechanisms can play an important role in ensuring that incentives for quality and unit cost containment are balanced. Despite their importance, research to place these plan provisions on a solid foundation of reason and empirical research is not well advanced.

The many challenges encountered in the design of benchmark incentive mechanisms for quality, combined with the dearth of good research in the field, make it reasonable to question whether such mechanisms are the best way to regulate quality in PBR plans. Continuation of traditional quality regulation, which holds the utility responsible for quality and obliges it to address any deficiencies, remains a sensible alternative. A hybrid system is also worthy of a consideration in which the utility is obligated to make regular reports on a set of quality indicators.

6 BENEFIT SHARING PROVISIONS

6.1 Introduction

As I explained in Section 2, a well-designed PBR plan generates stronger performance incentives with fewer operating restrictions than COSR. Performance is expected to improve under such a plan, and utilities can earn more and their customers pay less – at the same time – than could be the case under cost of service regulation. The details of a PBR plan will influence the allocation of plan benefits between utilities and their customers, and the proper mechanism for sharing plan benefits is a controversial issue in many PBR proceedings.

Benefit-sharing provisions should allow *both* shareholders and customers to fare better than under standard rate regulation. If PBR is voluntary, utilities have little incentive to agree to a plan unless it offers a reasonable chance for higher earnings, especially in view of the higher risk entailed. It is incorrect, then, to point to higher utility earnings under PBR as evidence of its “failure.” Higher utility earnings are consistent with successful PBR as long as customers also benefit compared with a continuation of the status quo.

The selection of a benefit sharing mechanism should be based on sensible criteria. I evaluate alternative sharing mechanisms primarily in terms of their effect in three areas: performance incentives, cross-subsidization, and risk reduction. Other attributes considered include simplicity and “salability,” (*i.e*., the ability to convincingly demonstrate benefit sharing).

Various PBR plan provisions influence customer benefits. These can be grouped into two general categories.[76] One is predetermined sharing provisions such as initial rate cuts and enhanced rate trajectory. These are so called because they are determined in advance of plan operation and are delivered to customers whether or not performance actually improves. A second general category of benefit sharing provisions is “real time provisions.” These include earnings sharing and cost-based rate resets.

In this section, I analyze the salient benefit-sharing provisions. I describe the basic features of each approach, detail important precedents, and evaluate its advantages and disadvantages as a means of benefit-sharing.

---

[76] Customer welfare also depends, of course, on the market responsiveness of rate and service offerings and service quality

6.2 Enhanced Rate Trajectory

One way to share the benefits of PBR is to enhance the rate trajectory so that it is more favorable to customers. Consider first how this might be done in the context of a rate or revenue requirement index. As we have already seen, the X-factor in such indexes influences allowed rate escalation. A higher value for X therefore benefits customers of regulated services. An X-factor designed in accordance with North American principles is calibrated to reflect the TFP trend of the relevant industry. One way to share expected plan benefits with customers, then, is to set the X-factor at a level above the calibration point. This component of the X-factor has been called, variously, a “consumer dividend” or “stretch factor.” It is set in advance to help ensure an external character for X. However, it can be allowed to vary from year to year.

The growth trend in rates is not the only way that customer welfare is affected by the rate trajectory. Customers are also affected by the extent to which the company absorbs operating risk. The base productivity factor, for example, is more than just the offering of the benefit of normal productivity growth. It is, furthermore, a commitment by the company to provide said benefit over a multi-year period during which actual productivity growth may be quite different. A rate freeze offers the customer protection against input price as well as productivity volatility.

An important advantage of stretch factors is that their values can be assigned independently of a company’s activities during the plan. Stretch factors therefore do not compromise performance incentives or operating flexibility. Valuations made prior to the first indexing period clearly have this attribute.

Critics of stretch factors have argued that regulators cannot commit to a stretch factor policy for subsequent plans. Absent such commitments, parties might reasonably expect stretch factors in future plans to reflect the utility’s productivity growth during the current plan. However, the brief history of US price cap regulation does not provide much evidence to support a concern about such stretch factor “ratcheting.”

To the extent that they are external, stretch factors are plainly not useful in reducing business risk. For example, the application of a stretch factor may give customers a 0.5% break in rates even if the company’s earnings are depressed by mild weather and a regional recession. As for regulatory risk, the short history of rate cap regulation provides few clear lessons. Critics of stretch factors argue that they lack the solid foundation in economic research that unit cost calibration points have. Regulators’ ability to assign values for stretch factors arbitrarily exacerbates the risk. On the other hand, the range of explicit stretch factor values that have been approved is actually fairly narrow. Virtually all have fallen in the 0 to 1.0% range.

The appropriate stretch factor depends in part on the prospects for productivity growth during the plan term. Expected productivity growth should by this logic be lower the greater is the efficiency of the company. Benchmarking studies can shed light on a company's operating efficiency. However, such studies invite controversy and good studies are expensive. Absent such work, regulators should take careful note of the regulatory system under which a company has operated. For example, efficiency should be greater for a company that has operated for many years without a rate case.

The incentives generated by the prospective PBR plan should also be considered in stretch factor selection. A plan that generates strong incentives should stimulate better performance, making more available to share. Analogously, a plan with weak incentives should have lower stretch factor. At the extreme, a plan for an efficient company that is expected to generate weaker incentives than COSR may in principle have a negative stretch factor. The extent of operating flexibility that a plan provides for also affects expected productivity growth. Marketing flexibility is especially important in this regard. Our reasoning then leads to the conclusion that a plan without much marketing flexibility should have a lower stretch factor.

Regarding their salability, stretch factors are appealing to regulators insofar as they represent an advance commitment to customer benefits. Customers therefore benefit whether or not performance improvements are realized—at least during the term of the plan. On the other hand, customers and their representatives may not understand that stretch factors are designed to be insensitive to a utility's current earnings and may resent high earnings if they occur. It is helpful in this regard for regulators to acknowledge the value of stretch factors and the long run benefits of high earnings when approving PBR plans.

6.3 Initial Rate Cuts

A less common approach to sharing plan benefits is to lower the initial (base year) rates or revenue requirement below the levels that would otherwise result. When this is done, consumers immediately reap a plan benefit. Moreover, benefits continue to be created in subsequent years since, with lower initial rates, lower prices result from index-based rate adjustments. This approach has been more widely used in Great Britain than in North American PBR to date.

The advantages and disadvantages of initial rate cuts as a benefit sharing mechanism are similar to those for stretch factors. To the extent that rate cuts do not deepen in successive plans in response to performance improvements, performance incentives are strong. Cuts at the outset of the first plan do not affect incentives. The concern is, instead, with the size of initial rate cuts that might occur at the start of *subsequent* plans and their linkage to past performance improvements under PBR. As with stretch factors, initial rate cuts do not mitigate business risk and can actually increase regulatory risk absent a proper conceptual and empirical foundation. Customers benefit whether or not utility performance improves but may resent high earnings if they occur.

A unique advantage of initial rate adjustments is the immediacy of the benefits. On the other hand, a unique disadvantage is the difficulty of demonstrating that rate cuts are in fact being made when, as is common, companies propose rate increases just prior to indexing. Utilities are then in the awkward position of claiming that they could have asked for even larger price increases and that customers have benefited from the company's restraint. Since other parties will have differing opinions about the warranted rate hike, the benefits may be less convincing.

Regulators considering initial rate cuts should recognize that they are in lieu of other benefit sharing provisions. For example, any initial rate cut should in principal reduce the appropriate stretch factor. This principal is clearly recognized in British-style PBR. Regulators in Britain and Australia explicitly discuss how plan benefits are to be divided between rate cuts and higher X-factors.

6.4 Earnings-Sharing

6.4.1 Description

An earnings-sharing mechanism (ESM) adjusts a company's price restrictions when its rate of return (ROR) has been in a certain range over a recent historical period. A typical ESM provides for rate adjustments when the actual (pre-sharing) ROR differs from a target ROR by certain prescribed amounts. The mechanisms are established in advance of their use and typically function for several years. The most widely-used rate of return in ESMs is return on equity (ROE).

Approved ESMs vary significantly in several ways. The most important difference is the shares of surplus (and/or deficit) earnings assigned to shareholders and customers. These shares may differ in different ranges around the target ROE. Many plans feature a deadband around the target in which rates are insensitive to ROE fluctuations. Immediately beyond the deadband, the customer share is commonly 50%. In some plans, it increases substantially when ROE is extraordinarily high and falls substantially when it is extraordinarily low. Thus, the company share falls with the extent of surplus earnings. ESMs with this attribute are sometimes called "regressive." Alternatively, a "progressive" ESM increases the company's share of benefits as surplus earnings increase.

Some plans are symmetric in the sense that they provide for rate decreases when earnings are high and similar rate increases when earnings are commensurately low. Other plans provide for rate adjustments only when earnings are high or low. For example, a plan approved for a Maine utility shares earnings deficits but not surpluses. Other plans share only surpluses. The symmetry of an ESM can, naturally, have a major impact on the risk-return balance of a PBR plan.

Precedents

United States

ESMs are one of the oldest approaches to PBR. They were used in England as early as 1855 to regulate local gas companies.[77] A plan was adopted in Canada in 1877 to regulate Consumers Gas. An early American plan was established in 1905 for the Boston Consolidated Gas. A plan for the Potomac Electric Power, approved in 1925, remained in effect until 1955. ESMs have been used recently by many US energy utilities. Most recent PBR plans for US and Canadian energy utilities involve ESMs. However, ESMs were not included in the PBR plans for National Grid (MA) or the plans approved by the FERC for oil pipelines or the power transmission services of International Transmission.

Experience with ESMs in the North American telecommunications industry is also interesting. Most of the early price cap plans at both the federal and state level included an earnings sharing mechanism (ESM) as an adjunct to the price cap mechanism. For example, the original FCC plan for the LECs included an ESM to provide a "backstop" in the event that the X-factors established by the FCC were substantially in error or in the event that a particular LEC's productivity significantly differed from the average.[78] In addition, the first price cap plans in California (Pacific Bell and GTE-California in 1990), New York (Rochester Telephone in 1991), Rhode Island (1992), and New Jersey (1993) all featured ESMs.

However, the FCC's later LEC price cap plan, adopted in 1997, did not include earnings sharing. The FCC believed that ESMs blunt the efficiency incentives created by price caps since companies must immediately share the benefits of efforts to reduce their unit costs.[79] The FCC also noted that "the removal of sharing also removes a major vestige of rate-of-return regulation that created incentives to shift costs between services to evade sharing in the interstate jurisdiction."[80] The FCC went on to state that the cost-shifting and cross subsidy incentives inherent in rate-of-return-based sharing mechanisms were at odds with the goal of promoting greater competition and eventually deregulating LECs, as envisioned by the Telecommunications Act of 1996:[81]

> Not only is sharing inconsistent with the general competitive paradigm that was established in the 1996 Act, but sharing might make it more difficult to deregulate services that become subject to substantial competition by creating an opportunity for LECs to misallocate costs from deregulated common carrier services to services that remain subject to sharing requirements. As more and more incumbent LEC services become subject to competitive pressures, the public interest detriments of the cross subsidy incentives inherent in sharing become worse as the costs that can be misallocated to services that remain subject to sharing requirements increase. Without the elimination of sharing, it might become necessary to adopt new structural or nonstructural safeguards to prevent or limit these misallocations. Rather than consider adopting such administratively burdensome requirements, I conclude that eliminating sharing is the more reasonable course.[82]

---

[77] For further discussion of the early precedents see Harry Trebing, "Toward An Incentive System of Regulation:, *Public Utilities Fortnightly*, July 18, 1963, p. 22-37.

[78] For example, see *Second Report and Order,* CC Docket 87-313, September 19, 1990, FCC 90-314, paras 120-165.

[79] For example, see *Fourth Report and Order,* CC Docket 94-1, May 7, 1997, FCC 97-159, para 148.

[80] Id

[81] Id, para 151.

[82] Some FCC Commissioners were even more adamant in their opinion about the negative features of earnings sharing. For example, Commission Chong stated that:

> "I am particularly pleased that this Report and Order puts a stake through the heart of 'sharing,' the requirement that incumbent LECs earning more than specified rates of return must 'share' half

Similarly, in state jurisdictions, ESMs are becoming increasingly rare as an adjunct to price cap plans. Few states currently use rate indexing in conjunction with an ESM to regulate the dominant LEC.

Canada

In Canada, ESMs have been fairly common in PBR plans for energy utilities. The OEB, for instance, approved the use of ESMs in plans for Union Gas and Ontario's power distributors. Several plans that lack ESMs have featured benchmark-style sharing mechanisms. Neither CRTC rate indexing plan for Canada's telecom utilities has featured ESMs.

Britain and Australia

Regulators in Britain have considered the adoption of ESMs on several occasions. One review of a British Gas plan featured an especially thorough deliberation of this issue. However, few ESMs have been adopted to date in Britain. A recent exception is the latest plan for the transmission system operation ("SO") services of National Grid. There are also no ESMs in the approved index plans for Australia's power transmission and distribution utilities.

6.4.2 Evaluation

ESMs have some important advantages as benefit sharing mechanisms. One is their ability to mitigate risk. This property is, of course, greater when ESMs are symmetric. ESMs are an automatic means of adjusting rates for a wide range of risky external developments. This can be appealing where risks are substantial or Commissions lack the technical expertise to approve alternative risk mitigation measures such as industry specific input price indexes. As an alternative to initial rate reductions and X-factor s, ESMs also reduce regulatory risk. In effect, benefits are shared *as realized* and there is less pressure on regulators to choose stretch factors and initial rate reductions that share the unknowable plan benefits. There is, however, some regulatory risk to the utility in proposing an ESM: principally, the risk that the Commission will approve an asymmetric ESM in which earnings shortfalls are not shared.

> or all of the amount above those rates of return with their access customers in the form of lower rates the following year. Since sharing continues the inefficiencies of a rate-of-return era, I have long believed that a system of pure price caps without sharing would be preferable. I believe that I have correctly found today that sharing tends to blunt the efficiency incentives I sought to create through the price cap plan."

Separate statement of Commissioner Rachelle B. Chong, *Fourth Report and Order*, CC Docket 94-1, May 7, 1997, FCC 97-159, p.2.

In addition to risk management, another benefit of ESMs is their popularity. Customers and their representatives appear to believe that ESMs align shareholder and customer interests. If a distributor had a 14% ROE last year, for instance, the ESM might reduce the revenue from regulated services by the value of 100 basis points of ROE. ESMs also help keep utility earnings within politically acceptable bounds.

On the downside, ESMs do not by themselves guarantee that customers benefit from a PBR plan. Customers may complain if distributor earnings exceed the target ROE but fail to reach the sharing range. Failure to reach the sharing range is especially likely when there are low initial rates or a high stretch factor. Customers must also remember that their rates may go up during an earnings shortfall.

Another disadvantage of ESMs is that the continued focus on earnings keeps alive inherently controversial issues like utility-affiliate transactions and cost allocations between a utility’s various regulated services and any competitive market services. This can give rise to controversies in ESM implementation hearings. Regulators may anticipate this and deny the company operating flexibility.

The effect of ESMs on performance incentives is complicated. Compared to a multiyear plan in which rate restrictions are completely insensitive to a utility’s performance, a plan with an ESM should in theory weaken performance incentives. After all, utility managers have less incentive to improve performance if half of the after-tax benefits go to customers. On the other hand, the practical reality is that the inclusion of an ESM in a plan may encourage interested parties to agree to an extension of the period between plan reviews. ESMs may also help the parties agree to plan termination provisions that have less deleterious incentive consequences. For example, it can be agreed that in the event of any cost based true-up of rates at the end of the plan, a company is entitled to keep its share of any surplus earnings and is not entitled to compensation for its share of surplus losses.

The analysis of the impact of ESMs on the direct cost of regulation has a similar flavor. ESMs increase regulatory costs during periods where companies are not otherwise subject to regulatory intervention, such as a multi-year rate plan. For example, with ESMs it may be necessary to compute the cost of regulated services, and therefore to allocate total cost between regulated and unregulated services.[83] This effect is offset to the extent that the inclusion of an ESM in a plan can persuade stakeholders to agree to extend the period between formal rate cases.

The reasons for the prevalence of ESMs in the approved PBR plans of North American energy utilities and their relative paucity in the PBR plans of telecom utilities merit brief consideration. Two explanations seem plausible. First, cost allocation issues have historically loomed larger for telecom companies than for energy utilities. Interstate access and local exchange services to business customers of LECs have long been subject to physical bypass, while residential customers have cellular bypass options and, increasingly, access to alternative land line providers. Because customers have so many alternatives to utility service, the marketing and cost allocation issues that result from ESMs may be more costly for telecom utilities.

A second reason for the discrepancy in the use of ESMs may be the relative novelty of PBR for energy utilities. As noted above, many early PBR plans for telcos featured ESMs, but earnings-sharing in the industry has become rarer over time. Similarly, ESMs may become less common for energy utilities as regulators and parties gain experience with PBR, including better knowledge as to all the costs associated with sharing mechanisms.

6.5 Plan Termination Provisions

Plan termination provisions are provisions for what happens to regulation on the occasion of a PBR plan's termination. These typically involve a formal rate case under both North American and British style index plan design methods. Two issues are salient in the specification of plan termination provisions. One is the plan term, which is the duration of time between formal rate cases. The other is the degree to which rate resets reflect other, external considerations.

6.5.1 Plan Term

Most PBR plans specify the term of their application. Formal rate cases will typically not be held during this term.

[83] This is a major concern for telecom utilities, which typically provide extensive regulated and unregulated services from the same facilities.

Precedents

The trend in PBR has clearly been towards plans of longer term. Plans of three year's duration were typical during the 1990's. More recently, five year terms have become standard and some plans of considerably longer duration have been approved. Especially noteworthy in this regard are the ten year plans for power distribution services of National Grid in Massachusetts and New York.

Evaluation

The rate case typically held at the termination of a plan is an important opportunity to share plan benefits with customers. Thus, short plan terms let customers share in benefits sooner. Short plan terms also reduce business and regulatory risk. This makes them more suitable for businesses undergoing rapid change or for regulatory jurisdictions where there is exceptional risk of unusual stretch factors or initial rate adjustments.

On the other hand, plans of longer duration strengthen performance incentives and alleviate concerns about cross-subsidies and novel operating practices that can lead to operating restrictions. Longer terms are especially useful in encouraging initiatives that involve up front costs to achieve long-run efficiency gains. That is one reason why longer plan terms are of interest in PBR plans occasioned by utility mergers. Both of the National Grid plans just mentioned involved mergers. The risk of a longer plan term can be reduced by several other plan provisions, including industry-specific inflation measures, Z-factors, marketing flexibility, and earnings sharing mechanisms.

6.5.2 Rate Reset Provisions

Description

The rate reset provisions of PBR plans can in principle involve widely varying degrees of externalization. At one extreme, rates may be reset entirely on the basis of a rate case and thus reset the company's rates to its cost and output. At the other, a plan could be reset entirely on the basis of external data. For example, a rate or revenue cap index could be revised only to better reflect the recent unit cost trend of the relevant industry.

The middle ground includes a number of possible options. One idea is to set the new rates as an average of the rates resulting from a new rate case and the rates resulting from one year's continuation of the old PBR mechanism. A related idea is to have as a revenue requirement a weighted average of the new cost of service as established in a rate case and of an external cost benchmark. If the company has been operating under an ESM, a third idea is to permit the company to keep its share of surplus earnings.

Precedents

Rate cases are a common input into the resetting of rates for energy utilities worldwide. However, plans do not always reset rates to exactly match a company's cost and output. For example, AmerenUE was permitted to keep some surplus earnings under an ESM at the time of a PBR plan update. External benchmarking is now used in extensively in several countries as an input to the rate reset process.

Several recently approved plans expressly provide for only a partial resetting of rates to reflect cost and output at the time of the plans termination. A salient North American example is the National Grid USA plan for power distribution in Massachusetts. In Britain and Australia, where rates reflect multi-year cost forecasts, several approved plans provide for companies to keep a share of lower than forecasted cost during the next plan.

Evaluation

Rate reset mechanisms have a major impact on customer benefits from PBR. A rate reset that is based entirely on a rate case passes to customers the full benefit of cost savings achieved. Risk is reduced.

Yet rate reset mechanisms also have a major impact on the incentives to make long term performance gains. To the extent that a full cost-based rate true-up is not ensured, performance incentives are strengthened and there are reduced concerns about cross subsidies and novel practices that can lead to operating restrictions. Incentives for initiatives involving up front costs and long term benefits are, once again, especially affected. Alternative mechanisms are available to mitigate risk that do not compromise performance incentives. As in the case of longer plan terms, a variety of other mechanisms are available to mitigate the resultant risk.

7 CONCLUDING REMARKS ON PBR PLAN DESIGN

Our discussion has revealed that many tools are available for the construction of PBR plans for energy utilities. These tools have differential impacts on performance incentives, operating flexibility and customer benefits. It is challenging to design a plan that strikes the right balance.

The benefits from PBR are maximized by plans that generate strong and balanced incentives for a wide array of activities. For example, plans should encourage utilities to strike the right balance of attention between cost containment and service quality. Benefits are typically greater for comprehensive rate or revenue cap plans than for non-comprehensive plans. Benefits are greater for price cap plans with marketing flexibility than for revenue caps, especially when they facilitate better utility marketing.

Our analysis has also highlighted the importance of encouraging energy utilities to undertake initiatives that involve up-front cost to achieve long term performance gains. Plan termination provisions play an especially critical role in the incentives for such initiatives. The greater risk of provisions that strengthen such incentives can be offset by more careful attention to eliminating unnecessary sources of operating risk under the plan.

Regarding the risk-return balance, careful plan design can help to achieve a risk-return balance that is right for utilities and their customers. Tools that reduce risk without unduly raising concerns about performance incentives and operating practices are especially desirable. For example, an industry-specific input price index can track fluctuations in a company's input prices better than a macroeconomic output price index. An X-factor based on a regional rather than a national TFP trend may better reflect the realistic expectation for unit cost growth. The Z-factor can reflect changes in government policy and other worrisome external developments.

The importance of tailoring plans to fit the circumstances of a utility must also be stressed. When it comes to PBR plan design, one size does not fit all. Utilities vary in their productivity growth expectations, risk exposure, and need for marketing flexibility. Different plans are therefore indicated if all are to properly balance risk, return, and customer benefit considerations.

8 APPLICATION TO ONTARIO POWER DISTRIBUTION

8.1 Rate Escalation Mechanism

8.1.1 British vs. North American Basis

Section 4 of our report discussed two fundamentally different approaches to the development of rate escalation indexes – the British and the North American. The latter approach focuses on industry input price and productivity trends. The power distribution industry of Ontario seems to be an excellent candidate for the North American approach given the nature of the business and the existence in the jurisdiction of a large number of similarly situated utilities. However, the British approach remains an option worth considering in other provincial energy businesses, including power generation and transmission

8.1.2 Inflation Measure

The candidate inflation measures for the rate escalation index include an industry-specific input price index and two macroeconomic price indexes: the GDPPI and the CPI. Both of these indexes are available at both the national and the provincial level. The industry specific approach has the strong appeal of reducing risk without weakening performance incentives. However, the design of such an index involves complex issues of capital cost measurement that are best considered in the development of the third generation PBR plan. A macroeconomic index should do a satisfactory job of compensating provincial distributors for the input price inflation that they are likely to encounter in the next 1-3 years. Of the available indices, the GDPPI is preferable to the CPI. In addition to the precedent for its use in Ontario, this index pertains to a more relevant set of goods and services. Its lessened sensitivity to fluctuations in the prices of foodstuffs and energy also makes it more stable.

8.1.3 X Factor

The options for the design of an X factor are considerably more varied than those for the inflation measure. The X factor can in principle be based on a new study of the input price and productivity trends of Ontario power distributors. This study could feature the development of a new productivity index. Alternatively, it could involve estimated productivity trends based on econometric cost research. The latter approach has the advantage of permitting custom projections for individual utilities that reflect their operating scale, customer growth, and other local business conditions. Econometric projections of TFP growth could also be developed from econometric cost research based on U.S. data.

As for stretch factors, a typical stretch factor in a rate escalation mechanism for an energy utility was shown in Section 4 to be around 0.5%. The brief duration of the contemplated plans argues for relatively low stretch factors. Stretch factors may in principle vary with estimates of the known operating efficiency of utilities. However, the available comparators and cohorts study is of experimental character and may not provide a solid empirical foundation for stretch factor recommendations. A linkage of stretch factors to benchmarking research may make sense in a third generation plan if there is further refinement in benchmarking methods.

Given the brevity of the plan period and the limited time and money budgeted for plan development, a more sensible approach to X factor selection may be to rely on historical precedents. The precedents listed in Table 1 are especially relevant since these decisions were informed by research on input price and productivity trends. If the Board chooses a macroeconomic inflation measure such as the GDPPI, our analysis suggests that the X factor precedents corresponding to this measure are especially relevant. X factors in all of the approved indexes for energy utilities that featured macroeconomic inflation measures ranged from 0.41% to 2.57% and averaged 1.16%. X factors in the three approved indexes for power distributors that featured macroeconomic inflation measures ranged from 0.63% to 2.57% and averaged 1.56%. The low number in this range has the advantage of being based on the most recent input price and productivity research.

8.1.4 Z Factor

Z factors were noted above to be a standard feature of rate escalation indexes. They protect utilities against changes in business conditions that are not reflected in the other terms of the index. Common reasons for Z factor adjustments include force majeure events and changes in government policy.

The need for Z factoring in the present application is limited by the short terms of the plans and the fact that several kinds of cost which are sensitive to government policy will be addressed by separate rate elements that are not subject to indexing. On the other hand, Ontario power distributors are vulnerable to ice storms and perhaps other *force majeure* events and existing rates frequently do not finance adequate contingency funds for these events. Under the circumstances, the plan should contain some provision for rate adjustments for a limited set of events. As always, the appropriate adjustments should reflect prudent measures to prepare for and respond to the events. Regulatory cost can be contained by establishing relatively high thresholds for qualifying events.

8.2 Earnings Sharing

The Board has devoted some thought to the possibility of earnings sharing in a second generation plan. Earnings sharing mechanisms were noted above to significantly raise regulatory cost and can weaken performance incentives in plans of short duration. I would counsel the Board against the use of earnings sharing in the second generation plan. If earnings sharing is adopted I would argue for relatively wide deadbands in the earnings sharing mechanism.