5

ELUSIVE EFFICIENCY AND THE X-FACTOR IN INCENTIVE REGULATION: THE TÖRNQVIST V. DEA/MALMQUIST DISPUTE

Jeff D. Makholm

Introduction

Incentive-based regulation is practiced worldwide, and all applications of it require some form of efficiency or productivity measurement—the *X-factor*. Including this factor in a multi-year regulatory formula allows the formula to survive intact for several years, and this longer regulatory lag between tariff reviews strengthens the incentives on firm performance. The factor, an index number, is intended to permit prices to move between tariff reviews according to an objective and reliable pattern. Differing opinions have arisen, however, on which index number to use.

One index number, the Malmquist Index, has generated considerable interest in some regions (particularly in Australia and Europe) because of its ostensible ability, when used in conjunction with data envelopment analysis (DEA), to distinguish readily between technical change for an industry (which the X-factor is generally held to measure) and efficiency for a particular firm. However, the DEA/Malmquist procedure for separating individual firm efficiency from technical change is inherently unreliable for identifying how inefficient a firm is. Neither the quality of data for regulated firms, nor the essentially idiosyncratic nature of such firms, supports an analysis of the *level* of efficiency of individual utilities. To the extent that regulators attempt to use the DEA/Malmquist procedure to set tariffs to reflect "efficient firm" standards, they inject unsupportable subjectivity and an unreliable methodology into a tariff-making process. The only reliable alternative is to estimate the X-factor directly by measuring long-run rates of change in efficiency indices. The Törnqvist index is best suited to this process, but other similar indices offer similar results.

The X-Factor in the Theory of Price Cap Regulation

Incentive regulation allows automatic or formulaic adjustment to regulated prices between tariff cases. That is, the plan controls the rate of change of the regulated firm's tariffs by adjusting a price cap (or revenue cap) annually according to a predetermined formula. The purpose is to ensure that price changes reflect changing costs the same way as in competitive markets: (1) Changes in industry prices track changes in industry costs and (2) the changes in an individual firm's prices relative to its costs differ from an industry average if its productivity growth differs from the average productivity growth of its industry.[1] This difference between the rate of change in industry prices and in individual firm costs causes a variation in profits. This is the carrot or stick with which the competitive process rewards efficiency gains and punishes firms that are slow to innovate, to reduce costs, or to respond to consumer demands.

The Place of Incentive Regulation in Regulatory Economics

Incentive regulation has been a key part of utility regulation for over 25 years. In that time, many regulated companies in North America and virtually all newly privatized companies around the world embraced under a variety of labels some form of incentive regulation. Generally, incentive regulation plans are characterized by a definite plan period, automatic adjustment for inflation, a productivity adjustment (the X-factor), and sometimes a way to share monetary gains between utilities and customers and/or reward (or penalize) quality of service changes. It is the X-factor that embodies the competition-like constraint to which regulated companies are held under incentive regulation. Imposing that constraint extends the period between tariff cases in an acceptable way and provides the time for cost-savings or sales maximizing incentives to pay off for investors. The X-factor is not an incentive in itself, but it permits regulatory formulae to stay in place longer—and that provides the incentive for more efficient long-term decisions on costs, sales, and investments.

In the early application of price cap regulation in the UK, a general notion existed that the X-factor was a variable simply subject to the regulator's choice. For example, Beesley and Littlechild describe the X-factor as "…a number specified by the government,"[2] as if it were some kind of bureaucratic target. More recent consensus is that the X-factor derives from a regulatory regime designed to limit monopoly utility prices over a defined number of years in a way that mimics the constraints that a competitive firm would face. In discussions on setting the appropriate X-Factor, economists generally agree with the theory set out above and on the two central elements of the relevant Total Factor Productivity (TFP) measures.[3] For example, Loube and Navarro confirm that a price cap plan begins with prices set so that the value of total inputs (including a normal return on capital) equals the value of total output for the company as well as the industry.[4] A number of writers confirm that the purpose of the price cap adjustment formula is to ensure that the constraint of regulated prices mimics the pressures that competition would place on a firm.[5] General agreement also exists among economists that the relevant TFP measure should be based on industry- rather than firm-specific productivity measures.[6]

Theoretical X-Factor Formulation[7]

The standard formulation for implementing price cap regulation is given by equation (5) from Appendix A:

$$dp = dp^N - X + Z \qquad (1)$$

where dp denotes a percentage growth rate in price, dp^N is the annual percentage change in a national index of output prices, and Z represents the change in unit costs due to external circumstances (which can be positive or negative).

If the industry achieves a productivity target of X and experiences exogenous cost changes given by Z, the price change that keeps earnings constant is given by equation (1). This price change is given by:

1. the rate of inflation of national output prices dp^N,
2. less a fixed productivity offset, the X-Factor, which represents a target productivity growth *differential* between the annual TFP growth of the industry and the whole economy,[8]
3. plus exogenous unit cost changes, written as the difference between the effects on the industry and economy-wide unit costs of the exogenous event.

To use the industry's productivity performance as a target for an individual company, rewrite equation (1) into the formula:

(2) $$PCI_t = PCI_{t-1} \times [1 + GDP - PI_t - X \pm Z_t],$$

where PCI_t is the value of the price cap index in year t, Z_t is the difference in the effects of exogenous changes on a specific company and on the rest of the economy, and $GDP-PI$ is the national output price index (i.e., "gross domestic product price index").

Simply put, the effect of using the above formula to limit price increases is that earnings remain the same if a company's achieved productivity differential just meets the target X-Factor. Thus a company must perform as well against economy-wide average TFP growth today as the industry as a whole has historically performed in comparison with economy-wide average TFP growth. If a company's productivity growth falls short of the target, its earnings will fall; if it exceeds the target, its earnings will rise. The price adjustment formula that sets this target adjusts output prices by: (1) the change in a national index of output prices less (2) the TFP growth target, measured as the difference between the change in industry TFP and that of the nation as a whole, plus[9] (3) the difference between the effect of exogenous changes on a company's costs and on the costs of the nation as a whole.

Thus, the historical relative TFP growth of the industry and the whole economy is taken as the target for the firm's TFP growth relative to the whole economy. National output price growth and exogenous cost changes are measured annually, but the X-Factor is fixed as the target amount by which TFP growth should exceed historical economy-wide TFP growth. This system of rewards and punishments sets up the same incentives as an unregulated firm would face in a competitive market, where failure to match industry average productivity growth results in lower earnings, and exceeding industry average productivity growth leads to increased earnings.

When turning to the empirical measurement of TFP, it is important to keep two points in mind: (1) the only relevant productivity measure is the *change* in TFP, not the *level* of TFP (discussed in Appendix A); and (2) it is only the *industry average* TFP growth that mimics the constraints faced by firms in a competitive market.

"X-Factor Quantification" and Index Numbers

This X-Factor lies at the heart of the discussion regarding the possible use of the DEA/Malmquist index to regulate utility prices as a component of price cap regulation. The X-Factor is ultimately an index number. Index numbers are found throughout the economy, expressing the value of some entity, like prices or gross national product, at a given period of time and in absolute number form, but related to some base period. Objectively determined incentive regulation uses such index numbers as the X-factor to reflect industry productivity growth.

The first issue concerning the empirical foundation of the X-factor is the use of long historical time trends in its calculations. The conventional assumption among productivity analysts is that the industry productivity and input prices are characterized by a valid and stable trend. This basic view of long-term trends has been adopted by many academic researchers who have studied macroeconomic time series such as GNP, prices, wages, unemployment rates, money stock, interest rates, etc. The issue of whether "structural breaks" disrupt such long-term trends has attracted considerable academic interest,[10] but it would appear that the stable trend hypothesis is a strong one and is most consistent with the search for objectivity in the calculation of a suitable X-factor. Using the longest historical data series consistent with available data allows analysts to identify the magnitude of the trend most reliably.

Since price cap regulation was introduced in the UK in the 1980s, and subsequently in the US in the early 1990s, considerable discussion has attended the choice of the index number to mimic productivity. Most of the literature on index numbers for productivity measurement pre-dates the use of such information in incentive regulation plans. Indeed, all three of the productivity index numbers in general use for price cap regimes were formulated by their named authors decades ago. They are the Fisher Ideal index, used by the Federal Communications Commission (the FCC) for telecommunications incentive regulation in the United States, the Törnqvist[11] index, which forms the basis for many electric utility TFP studies, and the Malmquist index, to which regulators in the Netherlands and Germany have referred on occasion (albeit for a different reasons).

Comparing the Törnqvist with Malmquist Indexes

The popularity of the Törnqvist index follows from its association with "translog" production and cost functions. Simply put, translog functions (which are functions squared in logarithms) were the first to allow economists to study empirically the "U-shaped" cost curves of real-life firms. With such functions, scale and substitution economies could be investigated empirically rather than assumed theoretically. With such flexible, empirically developed models of production technology as a foundation, the theoretical base for index numbers that reflect such production technology is very strong.[12] The translog multilateral productivity index[13] forms the basis for modern TFP studies in the electric power industry, including NERA's.

The Malmquist index in modern regulatory literature is usually mentioned alongside the Törnqvist index in the literature on index number theory. The two indexes are indeed close theoretical cousins. For regulatory purposes, however, various analysts have seized upon a particular feature of the Malmquist index that the Törnqvist does not share: the purported ability to measure the extent of inefficiency of individual utilities against supposedly more efficient peers. However, the use of DEA procedures along with the Malmquist index for the purpose of assessing individual firm efficiencies is not based on index number theory, nor is it consistent with the empirical applications for which it appeared in the literature. In this section I review the use of the Malmquist index by academic efficiency analysts as well as by index number theorists. I show that the use of that index in conjunction with DEA analyses to judge the efficiency of individual utilities is a particular misuse of an index number method, for which no support appears in the theoretical or empirical academic economic literature.

The Malmquist index arose in productivity theory as a more general, less restrictive, way of representing how a production function moves over time. Although it lends itself to the practice, it was not intended as a tool to "differentiate between technical change and changes in productivity."[14] It is not a use for which index number theorists investigated the Malmquist index nor is it supported in that literature.

In general, the Malmquist index measures the change in an industry's total factor productivity over time. It accounts for the fact that technology (i.e., best practice) is continually changing and that a firm's efficiency performance (relative to best practice) is also subject to change. For this reason, calculating this index requires a panel of data for the identification of both technological change and variations in firm efficiency. The Malmquist index describes productivity growth in terms of two components: (1) movements in the best practice frontier (i.e., technological change) and (2) shifts in firm efficiency that narrow or widen the gap between actual and frontier performance.

In comparison, the Törnqvist index does not decompose productivity growth in terms of technological change and efficiency "catch up," but rather in terms of the respective contributions of output and input growth (and their individual components if there is more than one) to the final result. Another important difference between these two estimation methods is that the Törnqvist index relies on cost shares or other value-based weights, which implies the use of price information in addition to quantity series, whereas the Malmquist index only requires quantity indexes to calculate productivity. Other than these differences, and provided that adequate data are available, the Törnqvist and Malmquist indexes should provide similar overall results for industry TFP.

The problem with the use of the Malmquist index is that it enables analysts to make assertions about firm-specific efficiency relating to its two components—one representing the "technology" and the other representing the "firm." The existence of the two components has led analysts to draw conclusions about the efficiency of a particular firm with respect to an industry standard—something that incentive regulation does not call for and that the quality of data to investigate the X-Factor does not support.

Data Envelopment Analysis (DEA) and the X-Factor

DEA combines multiple input and output measures (both monetary and physical) to generate an overall efficiency measure for a company. Mathematical programming methods allow researchers to apply quantitative information of a company and its peer group (i.e., the comparators) to determine relative efficiency performance.

Figure 1 illustrates the basic DEA approach. This figure displays an input-oriented[15] efficiency measurement for a group of 10 companies, which assumes that there is one type of output (e.g., MBTUs delivered) and two kinds of input (e.g., capital and labor). This type of efficiency measure considers the degree to which input quantities can be proportionally reduced without changing the output quantities. The figure plots the combination of inputs (x_1 and x_2) that each company employs to produce a unit of output, which for simplicity is normalized equal to one. Based on the actual behavior of the 10 companies, an envelope curve or efficiency frontier (shown in the Figure) is identified, reflecting the industry best practice. If the production function (which in this case has only two inputs) were to capture all the relevant determinants of cost, then the closer a firm is located to this curve the higher is its level of efficiency. In principle, firms that are located further out can produce the same amount of output with fewer inputs, bringing them closer to the origin and the achievement of higher efficiency. Theoretically, each firm's efficiency level can be measured empirically. For instance, Firm P's score

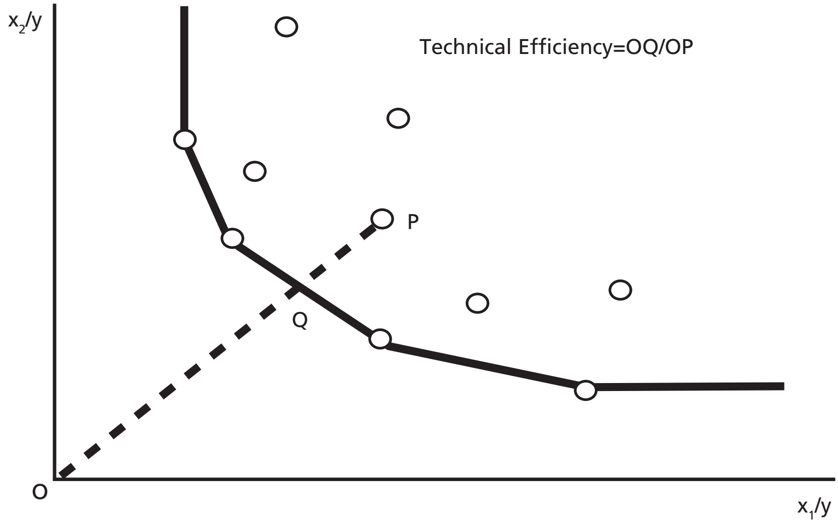


Figure 1. Efficiency Measurement with Data Envelopment Analysis (DEA)

is equal to the ratio OQ/OP. If a firm is located on the frontier, then it obtains the highest possible score, which is equal to one.

Certain analysts (and some regulators) have taken the relative positions on such graphs as Figure 1 as indicating what the X-Factor should be for a particular firm, for instance, by calculating an "efficiency score" for each company equal to the distance from the "efficiency" line. However, these conclusions are inconsistent with the price cap theory that uses a competitive type of constraint for multiyear regulated prices precisely because such conclusions ignore the fact that relative productivity levels are elusive when particular utilities are highly idiosyncratic. Any conclusions about relative efficiency are limited by the caveat that the DEA analysis measures all relative cost drivers. In practice, for utilities in different locations, with different histories, serving different kinds of customers, this is quite obviously not the case. That is, while such an analysis can be useful in gauging the relative efficiency in very similar operations (like McDonald's franchises, which operate from similar shops selling similar, or even identical, products), the same is definitely not true for different utilities selling to different customer bases in different regions of a country (or the world). In such cases, the gap between the company and the frontier could as well be due to any factors not recognized in the analysis and is not necessarily a measure of "inefficiency levels" or "productivity levels."

The DEA/Malmquist Procedure in Efficiency Analyses

Users of the Malmquist index number in regulatory settings frequently refer to the "seminal" 1978 paper by Charnes, Cooper, and Rhodes.[16] This paper is about measuring efficiency "with special reference to possible use in evaluating public programs."[17] In that paper, Charnes, et al. use DEA as a method to chart the comparative efficiency of public programs (decision making units—DMUs). That analysis (the graphical representation is shown above in Figure 1) measures the distance between the presumed efficiency frontier and the position of an individual DMU, implying inefficiency in that unit. They do, however, warn of the method's limitations outside of the public setting, saying

> One limitation may arise because of lack of data availability at individual [decision making unit] levels. This is likely to be less of a problem in public sector, as contrasted with private sector, applications. ... Our measure is intended to evaluate the accomplishments, or resource conservation possibilities, for every DMU with the resources assigned to it.[18]

By acknowledging the need to standardize the "resources assigned to it," as in the case of their school district example, the authors recognize the limitations of their suggested DEA method in situations where input choice or environmental factors cannot be controlled. Despite its limitations for private firms, DEA analysis is a direct analog to the Malmquist index, where the "distance" of a particular firm's observation (in a particular year or for an average of years) is compared to the "envelope." Like DEA analysis generally, the most fundamental problem with using the Malmquist index in this way for different network utilities is that neither all the input choices nor all the environmental factors can be controlled. Individual regulated firms exist in specific local surroundings. The myriad important factors (age, location, vintage of capital stock, idiosyncratic local regulation, etc.) create cost or output differences for particular utilities that their regulatory data does not (and can never hope to) capture. This type of comparison confuses these ubiquitous differences in conditions for significant differences in efficiency.

Federico went right to the heart of the problem of ignoring variations in environment issues:

> In spite of its nice theoretical properties, the Malmquist index is subject to all the shortcomings of conventional measures. It does not take into account environmental [factors], nor possible distortions from the use of benchmark years and the two measures of technical change differ if technical progress on the "frontier" is not neutral. On top of this, the Malmquist index (as the multi-country production function estimates) assumes that all units can attain the same level of production given their factor endowment—i.e., that they belong to the same production function. This assumption may not hold in agriculture, where feasible techniques heavily depend on environment.[19]

What is true of agriculture is true of any business—including network utilities—where local conditions dictate the precise form of investments and operations. The question of environmental factors cannot be disentangled from efficiency in either DEA analysis or its Malmquist equivalent. Sena reviews the various methods and warns about these environmental variables in evaluating the results of either DEA or Malmquist models that purport to identify efficiency for individual not on the frontier:

> However, the main weakness of DEA (namely that it is a deterministic method) is still there and so the computed distance functions may include the effect of factors not related to technical efficiency and technical change. ... The best option left to the researcher is to try to specify the DEA model (underlying the Malmquist index) in the best possible [way]... to minimize the impact of external factors on the computed distance functions.[20]

Sena also identifies another problem with the use of DEA analyses underlying the Malmquist index—that of stochastic shocks in the data:

> DEA does not allow us to model stochastic shocks to production i.e., it is deterministic. Therefore the computed efficiency scores may be biased by factors which are external to the production process. Not surprisingly, some attempts have been made to incorporate stochastic components into the linear programming problem. ...The

> data requirements of the chance-constrained efficiency measurement, however, are too many. Indeed it is necessary to have information on the expected values of all variables, along with their variance and covariance matrices and the probability levels at which feasibility constraints are to be satisfied. Therefore, this approach is too informationally demanding to be implemented easily.[21]

The issues associated with bias due to stochastic shocks are genuine and highly problematic for DEA analyses with electric utility data. Appendix C to this paper contains TFP data computed for a 1986 study of electric utilities,[22] using Form 1 data from the Federal Energy Regulatory Commission (the FERC) using the Uniform System of Accounts.[23] The productivity growth figures displayed in the Appendix, generated with a Törnqvist aggregation using the most reliable and consistent data for 39 electric utilities across 11 years, still shows considerable levels of stochastic shocks, particularly in year-to-year comparisons. For example, Kentucky Power for the four years 1973 through 1976 shows TFP yearly growth rates of −22.4 percent, 20.6 percent, −20.2 percent and 28.1 percent. The average TFP growth for Kentucky Power for the 11 years is 3.2 percent, and for those four particular years is 1.6 percent. But a DEA analysis of cost levels in 1974 or 1976 would incorporate very high productivity growth—owing only to stochastic shocks that were reversed in the next year—and those numbers make other companies in those years seem less productive by comparison.

Empirical data from academic TFP studies show that even the highest quality data (from the U.S. Uniform System of Accounts) produces TFP index growth rates for individual companies that are highly sensitive to vagaries and judgments on how company data is reported to governmental agencies. Individual data points for specific companies and years in industry-wide TFP analysis are notoriously unstable, even in the best of circumstances (see the data in Appendix C). The DEA envelope process, or the Malmquist index method, necessarily picks up the instability in individual data points and represents a stochastic error as a shift in technology. Simple noise among a cross-section would be taken as a change in the frontier—an advance of productivity. The more "noise" there is in the data, the more it pushes the envelope, implying inefficiency where none would otherwise be shown to exist. Thus, a simple DEA Malmquist analysis would treat the advances of companies in panel data TFP analyses as a shift in technology and would consider retreats as inefficiency.

In any event, to the extent that particular firms enter and leave the technology envelope on a short-term basis (which is indeed the case with the TFP data I analyzed in Appendix C), that envelope has no reliable significance as an indication of technological possibilities. Given that the envelope encapsulates unreliable individual data points and overstated technical progress, any conclusions based on the technological change and the efficiency "catch up" components of the Malmquist index would be highly unreliable.

Nevertheless, jurisdictions continue to rely on the Malmquist index in their DEA analyses. The Independent Pricing and Regulatory Tribunal (IPART) in New South Wales, Australia, has commissioned a number of regulatory benchmarking studies using the DEA/Malmquist technique.[24] These studies measure DEA production frontiers as a yardstick against which to measure the relative performance of the distributors under IPART's jurisdiction. Recent analyses have also been performed comparing the efficiency of individual Dutch electricity generators.[25] Another analysis was performed for German electricity distributors in the Federal network regulator's (BNA's) 2006 report on incentive regulation.[26] Scandinavian regulators routinely use such studies. These regulatory applications reflect a similar use of the DEA/Malmquist technique, with a similar justification:

> The Malmquist index ... can be decomposed so that the change in total factor productivity may be separated into a shift of the frontier (technical change) and a shift relative to the frontier (change in efficiency).[27]

This reasonable-sounding goal is contrary to the role of productivity in the theory of incentive regulation, as outlined in Section II and Appendix A, and, even if this were a valid pursuit in incentive regulation, it is contrary the advice of Federico and Sena regarding the difficulty of standardizing environmental factors. DEA's adherents seem to like the ease with which it provides "efficiency scores" for particular utilities. But that ease of calculation both contradicts the theory upon which incentive regulation rests and remains inconsistent with the kind of data available for utilities to which DEA is applied.

Summary of the DEA/Malmquist Procedure

Given the characteristics listed above of the Malmquist index and of DEA, any plan to base a price cap on the separation of technological change from company efficiency is going to run into problems than cannot be overcome in an objective manner. The DEA/Malmquist procedure cannot possibly control for all the environmental factors that determine a company's performance. Moreover, random shocks ("noise") in these unexplained factors can lead to further downward bias in the "frontier" and hence to a further underestimate of a company's performance.

The X-Factor remains a highly useful part of incentive regulation. The DEA/Malmquist procedure, however, is a devilishly convenient but ultimately unreliable procedure, inconsistent with the principles of incentive regulation. It is based on assumptions of production technologies and not on theory supported by the economic literature or valid empirical work. It has no support in the economic literature on the theory of index numbers and is contrary to the accepted theory regarding the incentives that price caps are supposed to embody. It is also contrary to the use of the DEA/Malmquist procedure in the analysis of nonregulated businesses where in contrast to network operations the inputs are controlled, and it has manifestly clear and unavoidable empirical problems.

Appendix A

The Derivation of the PBR formula:

Assume the price cap plan begins with appropriate prices so that the value of total inputs (including a normal return on capital) equals the value of total output for the company as well as the industry. For the industry, we can write this relationship as

$$\sum_{i=1}^{N} p_i Q_i = \sum_{j=1}^{M} w_j R_j$$

where the industry has N outputs ($Q_i, i=1,\ldots,N$) and M inputs ($R_j, j=1,\ldots,M$) and where p_i and w_j denote output and input prices, respectively. We want to calculate a productivity target for a company based on industry average productivity growth.

Differentiating this identity with respect to time yields

$$\sum_{i=1}^{N} \dot{p}_i Q_i + \sum_{i=1}^{N} p_i \dot{Q}_i = \sum_{j=1}^{M} \dot{w}_j R_j = \sum_{j=1}^{M} w_j \dot{R}_j$$

where a dot (•) indicates a derivative with respect to time. Dividing both sides of the equation by the value of output ($Rev = \sum_i p_i Q_i$ or $C = \sum_j w_j R_j$), we obtain

$$\sum \dot{p}_i \left(\frac{Q_i}{REV}\right) + \sum \dot{Q}_i \left(\frac{p_i}{REV}\right) = \sum \dot{w}_j \left(\frac{R_j}{C}\right) + \sum \dot{R}_j \left(\frac{w_j}{C}\right)$$

where *REV* and *C* denote revenue and cost. If rev_i denotes the revenue share of output *i*, and c_j denotes the cost share of input *j*, then

(1) $$\sum_i rev_i\, dp_i = \sum_j c_j\, dw_j - \left[\sum_i rev_i\, dQ_i - \sum_j c_j\, dR_j\right]$$

where *d* denotes a percentage growth rate: $dpi = \dot{p}_i/p_i$. The first term in equation (1) is the revenue-weighted average of the rates of growth of output prices, and the second is the cost-weighted average of the rates of growth of input prices. The term in brackets is the difference between weighted averages of the rates of growth of outputs and inputs. It thus is a measure of the change in TFP. Rewriting the equation for clarity, we see that

$$dp = dw - dTFP.$$

In other words, the theory underlying the annual price cap adjustment formula implies that the rate of growth of a revenue-weighted output price index is equal to the rate of growth of an expenditure-weighted input price index plus the change in total factor productivity (TFP). This equation shows that TFP is the appropriate foundation for a productivity target in the price cap plan: If the price cap plan begins with revenues that just match costs for a company, and if it attains the same productivity growth as the industry (measured in terms of TFP), then that company's revenues will continue to match its costs.[28]

Applying this rule more generally to admit the possibility of exogenous cost events outside of a regulated company's control, we may write

$$dp^* = dw - dTFP$$

where dp^* represents the annual percentage change in industry output prices inclusive of these exogenous costs, and dw represents the annual percentage change in input prices. To raise or lower industry output prices in order to track exogenous changes in cost, we write

(2) $$dp = dw - dTFP + Z^*$$

where *dp* represents the annual percentage change in industry output prices adjusted for exogenous cost changes, and Z^* represents the unit change in costs due to external circumstances.[29] Thus, to keep the revenues of the industry equal to its costs despite changes in input prices, the price cap formula should (1) increase industry output prices at the same rate as its input prices less the target change in productivity growth, and (2) directly pass through exogenous cost changes.

Equation (2) sets the allowed price change as input price changes less TFP growth adjusted for exogenous cost pass-throughs. If the economy-wide inflation rate were *assumed* to be the measure of the industry's input price growth and the X-Factor were similarly *assumed* to be its TFP growth target, equation (2) would indeed be the basis for the ideal price adjustment formula. However, these two assumptions are incorrect:

1. Broad inflation measures capture national *output* price growth, not the industry's input price growth. So even if the industry were a microcosm of the whole economy, a measure that captures national output price growth would not be an appropriate measure of its input price growth.[30]
2. The *X-Factor* is a target TFP growth rate relative to the economy as a whole (or relative to the TFP growth already embodied in national output price growth). The change in TFP in equation (2) is the absolute TFP growth for the industry. Again, unless economy-wide TFP growth is zero, the *X-Factor* is not equal to *dTFP*.

To get from equation (2) to the price adjustment formula, we must compare the productivity growth of the industry with the productivity growth of the whole economy. It is difficult to measure input price growth objectively. We are unaware of any agency that maintains an index of industry-specific input prices. Further, a productivity adjustment based on company-provided calculations of changes in their own input price index would be controversial and would not necessarily be based on information outside the company's control. However, by comparing productivity growth of the industry with that of the whole economy, we avoid the difficulty of measuring input price growth.

For the economy as a whole, the relationship among input prices, output prices, productivity, and exogenous cost changes can be derived in the same manner as it was derived in equation (2) above

(3) $$dp^N = dw^N - dTFP^N + Z^{*N}$$

where dp^N is the annual percentage change in a national index of output prices, dw^N is the annual percentage change in a national index of input prices, $dTFP^N$ is the annual change in the economy-wide total factor productivity, and Z^{*N} represents the change in national output prices caused by the exogenous factors included in equation (2). Subtracting equation (3) from equation (2) gives

$$dp - dp^N = [dw - dw^N] - [dTFP - dTFP^N] + [Z^* - Z^{*N}],$$

or

(4) $$dp = dp^N - [dTFP - dTFP^N + dw^N - dw] + [Z^* - Z^{*N}]$$

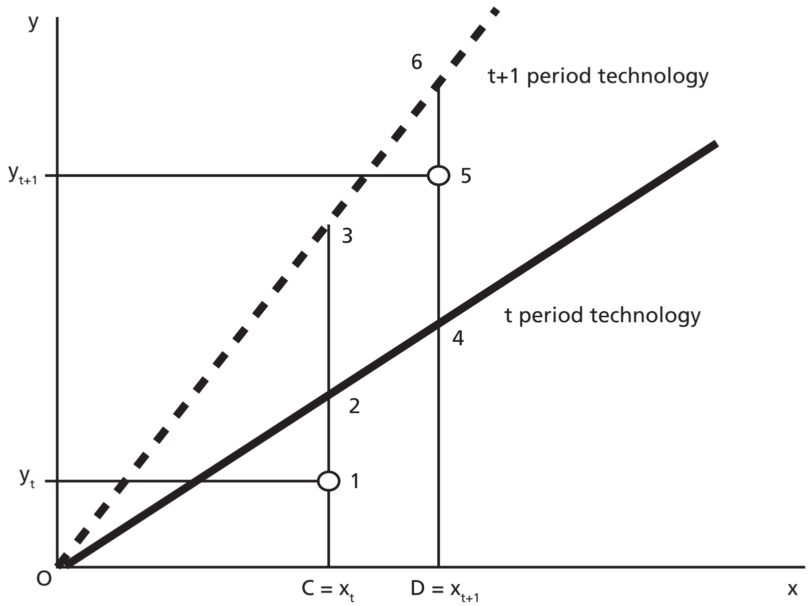
which simplifies to

(5) $$dp = dp^N - X + Z.$$

Appendix B

The Malmquist Index

Figure 2 illustrates the measurement of the Malmquist index, assuming an output-oriented efficiency measure and a constant return to scale technology. To simplify the exposition, I consider one output and only one type of input category. Figure 2 shows the efficiency frontier and a firm's output/input combination for two different time periods. Point 1 refers to initial period (time t), and point 5 pertains to the second period (time t+1). Based on the t-period technology, the firm's initial efficiency is measured by the distance C1/C2, and using the following period technology as reference, it is equivalent to the ratio C1/C3. A similar calculation is made regarding the firm's performance in the following period, so that based on the initial period technology its efficiency is measured as D5/D4, and with the t+1 technology, it is equal to the distance D5/D6. The Malmquist index combines productivity information relative to actual efficiency behavior and best practice frontiers in both periods in order to determine the efficiency change (or productivity growth) between the t and t+1.



$$\text{Malmquist Index} = \left(\frac{D5/D4}{C1/C2} \; \frac{D5/D6}{C1/C3} \right)^{1/2}$$

Figure 2. Output-Oriented Malmquist Index

Appendix C

YEARLY GROWTH RATES FOR
TFP INDEX
FOR 39 ELECTRIC UTILITIES

COMPANY	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	AVG.
POTOMAC ELECTRIC POWER	14.1%	5.5%	-7.2%	-5.4%	-3.6%	3.3%	-8.6%	8.3%	-5.5%	-0.6%	.0%
GULF POWER COMPANY	-0.1%	-8.4%	12.0%	-5.5%	-4.6%	-6.5%	-3.1%	3.5%	3.9%	-11.3%	-2.0%
TAMPA ELECTRIC COMPANY	0.6%	1.1%	3.5%	-7.5%	-5.2%	3.3%	6.2%	0.3%	0.2%	7.6%	1.0%
SAVANNAH ELEC AND PWR CO	-1.7%	-1.5%	1.8%	-12.9%	1.6%	4.8%	0.4%	-9.6%	1.4%	-13.8%	-2.9%
HAWAIIAN ELEC PWR CO	4.0%	2.2%	-0.3%	1.8%	2.9%	0.4%	1.7%	-0.4%	3.0%	-0.1%	1.5%
COMMONWEALTH EDISON	-8.8%	-8.4%	-1.9%	-2.0%	-15.6%	-6.1%	0.9%	-5.7%	-4.4%	-6.2%	-5.8%
INDIANAPOLIS PWR AND LIGHT	-5.5%	-5.1%	3.8%	-1.1%	-6.6%	1.3%	-10.2%	12.8%	1.2%	3.0%	-0.6%
PUB SERV OF INDIANA	1.6%	6.0%	0.4%	-4.7%	2.8%	4.5%	-1.7%	-5.3%	3.3%	-3.3%	0.4%
KANSAS GAS AND ELECTRIC	-6.3%	-5.4%	-2.7%	5.4%	3.6%	-5.8%	10.0%	11.6%	-9.7%	4.6%	0.5%
KENTUCKY POWER COMPANY	6.7%	0.1%	-9.4%	6.0%	-10.2%	13.6%	-6.2%	-5.3%	-6.8%	13.2%	0.2%
KENTUCKY UTILITIES COMPANY	30.5%	-3.4%	4.2%	15.0%	-1.5%	5.3%	2.1%	0.6%	-13.6%	8.3%	4.8%
LOUISIANA PWR AND LIGHT	12.1%	6.3%	0.2%	-4.3%	2.5%	2.1%	-4.9%	-3.2%	-9.9%	-7.4%	-0.6%
DETROIT EDISON COMPANY	-0.7%	-1.8%	1.1%	-0.2%	-3.2%	-4.6%	-0.2%	-1.2%	-2.6%	-6.0%	-1.9%
MISSISSIPPI POWER CO	-6.9%	-7.5%	4.1%	3.1%	-9.6%	1.8%	-14.5%	13.0%	-8.2%	2.1%	-2.3%
MISSISSIPPI PWR AND LIGHT	-3.0%	7.1%	-25.7%	-3.5%	16.4%	15.5%	8.1%	5.0%	-9.7%	1.7%	1.2%
KANSAS CITY PWR AND LIGHT	-0.3%	-2.5%	-3.9%	-11.1%	-5.2%	-3.2%	-6.5%	4.0%	-30.0%	29.4%	-2.9%
UNION ELECTRIC COMPANY	1.8%	.0%	11.5%	-6.5%	10.2%	3.8%	6.4%	-2.2%	-4.3%	-5.0%	1.6%
NEVADA POWER COMPANY	10.8%	-4.1%	-0.7%	8.1%	-2.6%	10.6%	14.5%	-12.6%	0.4%	-1.6%	2.3%
PUB SERV OF NEW HAMPSHIRE	-7.0%	3.2%	-10.1%	-9.4%	2.2%	-2.3%	3.9%	-14.8%	14.6%	-2.1%	-2.2%
PUB SERV OF NEW MEXICO	5.0%	0.1%	-5.4%	2.6%	-0.7%	-16.8%	5.0%	-34.1%	-0.1%	1.1%	-4.3%
OTTER TAIL POWER CO	-8.6%	2.6%	8.6%	-6.3%	14.6%	8.3%	-0.8%	-8.1%	-11.5%	7.8%	0.7%
CLEVELAND ELEC ILLUM CO	8.0%	-1.0%	1.6%	-5.7%	-4.7%	-2.3%	1.4%	-3.5%	-10.2%	-7.1%	-2.4%
COLUMBUS AND SOUTHERN OHIO	1.1%	8.0%	6.2%	-2.8%	-7.8%	4.5%	-1.6%	-6.9%	15.5%	-6.8%	0.9%
OHIO EDISON COMPANY	-6.4%	6.7%	-0.4%	-23.3%	-8.6%	9.0%	-15.4%	10.1%	3.9%	.0%	-2.4%
OKLAHOMA GAS AND ELEC CO	1.0%	4.3%	5.9%	-1.1%	-2.0%	-4.4%	-4.8%	9.4%	-1.6%	0.7%	0.7%
PUB SERV CO OF OKLAHOMA	5.7%	0.2%	-5.6%	6.0%	0.3%	3.3%	-1.5%	-0.2%	-8.4%	4.3%	0.4%
DUQUESNE LIGHT COMPANY	0.9%	2.0%	2.9%	20.9%	-5.2%	5.7%	-1.3%	-17.5%	18.5%	1.2%	2.8%
PENNSYLVANIA PWR AND LIGHT	5.6%	13.7%	10.2%	-4.9%	6.4%	-1.8%	4.7%	-5.5%	-0.4%	-3.4%	2.5%
CENTRAL POWER AND LIGHT	9.2%	-5.1%	0.2%	-4.4%	-4.4%	0.9%	2.5%	4.0%	-3.5%	-3.5%	-0.4%
DALLAS POWER AND LIGHT CO	3.0%	3.4%	0.9%	4.1%	6.3%	4.2%	0.4%	1.2%	2.0%	3.4%	2.9%
EL PASO ELECTRIC CO	0.2%	5.2%	2.5%	0.6%	2.0%	-5.0%	0.3%	-9.5%	6.8%	-10.7%	-0.8%
HOUSTON LIGHTING AND PWR	1.2%	1.5%	-1.5%	-4.1%	-1.6%	1.3%	-2.6%	-3.3%	0.6%	-3.8%	-1.2%
SOUTHWESTERN ELEC PWR CO	0.3%	11.1%	-9.7%	3.3%	-1.3%	-4.0%	0.6%	0.5%	1.2%	-2.2%	.0%
SOUTHWESTERN PUB SERV CO	3.4%	5.2%	-0.9%	1.3%	-3.1%	2.8%	0.1%	-0.9%	-1.4%	5.0%	1.2%
TEXAS ELEC SERV CO	-1.3%	0.6%	-2.1%	5.1%	1.7%	1.6%	-1.3%	5.8%	2.9%	1.0%	1.4%
TEXAS PWR AND LIGHT CO	1.1%	-2.3%	-1.8%	-9.4%	-8.1%	-5.8%	6.0%	-0.8%	-5.8%	1.9%	-2.5%
WEST TEXAS UTILITIES CO	3.0%	5.3%	-2.0%	3.2%	3.9%	1.9%	1.7%	-4.1%	3.0%	2.5%	1.8%
UTAH PWR AND LIGHT CO	-13.4%	21.7%	21.0%	-4.8%	6.9%	-23.8%	31.7%	15.1%	-2.1%	14.6%	6.7%
APPALACHIAN PWR CO	10.5%	26.3%	-3.0%	-11.9%	-9.1%	4.8%	-6.2%	-1.3%	1.7%	1.2%	1.3%
AVERAGE	1.8%	2.4%	0.2%	-1.7%	-1.0%	0.7%	0.4%	-1.3%	-1.7%	0.5%	.0%

Notes

1. The theory of incentive regulation, as derived in Appendix A, deals with the constraints posed by productivity *growth*. The *level* of productivity, as such, is not a focus of the economic concepts that form the basis of incentive regulation.

2. M. Beesley and S. Littlechild, "The Regulation of Privatised Monopolies in the United Kingdom," *The Rand Journal of Economics,* XX, 3 (1989), p. 455; also see M. Armstrong, S. Cowan, and J. Vickers, *Regulatory Reform: Economic Analysis and British Experience* (Cambridge, MA and London: MIT press, 1994), p. 174 for a discussion on the flexibility available to regulators when setting the *X-factor*.

3. That is, (1) changes in industry prices track changes in industry costs and (2) the changes in an individual firm's prices relative to its costs differ from the industry average due to its relative TFP growth.

4. R. Loube, "Price Cap Regulation: Problems and Solutions," *Land Economics,* LXXI, 3 (1995) 288; and P. Navarro, "The Simple Analytics of Performance-Based Ratemaking: A Guide for the PBR Regulator," *Yale Journal on Regulation,*

XIII, 1, (1996) 128. For further discussions on the importance of the correct price level when setting X see J. Bernstein and D. Sappington, "Setting the X Factor in Price-Cap Regulation Plans," *Journal of Regulatory Economics,* XVI, 1, (July 1999) 9, 11; and I. Vogelsang, "Optimal Price Regulation for Natural and Legal Monopolies," *Economia Mexicana, Nueva Epoca,* VIII, 1 (1999) 31.

5. J. Bernstein and D. Sappington, "How to Determine the X in RPI-X regulation: A User's Guide," *Telecommunications Policy,* XXIV, 1, (2000) 64. For additional discussions on the intention to track efficient costs by X tracking the differences in input price and productivity growth rates between the relevant industry and the economy, see Vogelsang (1999) p. 10, Bernstein and Sappington (2000) page 64, J. Vickers and G. Yarrow, Privatization: *An Economic Analysis* (Cambridge, MA and London: MIT Press, 1989) p. 296; and Loube (1995), pp. 289-290.

6. See: Loube (1995), p. 289.

7. This theoretical presentation, derived in Appendix A, is taken from J.D. Makholm and M. J. Quinn, "Price Cap Plans for Electricity Distribution Companies Using TFP Analysis," NERA Working Paper (October 21, 1997) pp. 36-39.

8. This differential is equal to the difference between the electricity industry and economy-wide TFP growth rates only if the rates of input price growth are the same for the industry and the nation, i.e., if $dw = dw^N$.

9. Adjusted for observed differences between input price growth rates for the industry and the nation.

10. In an influential article, Charles Nelson and Charles Plosser postulate that macroeconomic variables are better characterized as "non-stationary" processes that have no tendency to return to a predetermined path, instead of being regarded as variables that fluctuate around a deterministic trend. See Charles R. Nelson and Charles I. Plosser, "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications," *Journal of Monetary Economics* X (1982), 139-162. Pierre Perron, on the other hand, makes one of the most compelling defenses of the "trend-stationary" model, arguing that the empirical evidence validates this model when one accounts for the existence of trend-breaks due to certain "structural shocks" that have lasting effects See Pierre Perron, "The Great Crash, The Oil Price Shock, and the Unit Root Hypothesis," *Econometrica,* LVII, 6 (1989), 1361-1401. Perron finds that the only shocks with persistent effects are the 1929 Great Crash and the 1973 oil price shock.

11. Törnqvist (a statistician in Finnish government service writing in the 1930s) and Theil (an American econometrician) both investigated the validity of index number techniques. The index number used most widely for TFP studies, which is the geometric mean of the Laspeyres and Paasche indexes described in basic economics textbooks, is named after both.

12. In technical terms, the Törnqvist/Theil index number is "exact" for the flexible homogeneous translog aggregator function. The Index is "exact" in the sense that it can be directly related to the properties of the translog. For further reference, see W. E. Diewert, "Exact and Superlative Index Numbers," *Journal of Econometrics,* IV, 2, (1976), 115-146.

13. D.W. Caves and L.R. Christensen, "Global Properties of Flexible Functional Forms," *American Economic Review,* LXX, (1980) 422-432.

14. M. Dykstra, "How Efficient is Dutch Electricity Generation: Current Research," CPB Report (the Netherlands), 1997/4, pp. 45-47 (http://www.cpb.nl/nl/pub/cpbreeksen/cpbreport/1997_4/s3.pdf)

15. DEA also allows the construction of output-oriented efficiency measures, which we describe later on with regard to the issue of total factor productivity. In this case, the relevant question is, by how much can output quantities be proportionally expanded without altering the input quantities used? Output- and input-oriented measures are equivalent only in those cases in which the technology of production exhibits constant returns to scale.

16. A. Charnes, W.W. Cooper and E. Rhodes, "Measuring the Efficiency of Decision Making Units," *European Journal of Operational Research,* II (1978), 429-444.

17. Ibid., p. 429.

18. Ibid., p. 443.

19. G Federico, "Why are we all alive? The Growth of Agricultural Productivity and its Causes, 1800-2000," European University Institute, paper for the Sixth conference of the European Historical Economics Society, Istanbul, 9-10 September 2005, pp. 4-5.

20. V. Sena, "The Frontier Approach to the Measurement of Productivity and Technical Efficiency," *Economic Issues,* VIII, Part 2 (2003), 90. Sena refers to the DEA model "underlying the Malmquist index" in the sense that the latter index is a specific application of the general "DEA model" approach to measuring distance between a particular observation and the frontier. She does not imply that the DEA model and the Malmquist index are anything more than analogues in this respect.

21. Ibid., p. 83.

22. The data in Appendix C appears in J.D Makholm, "Sources of Total Factor Productivity in the Electric Utility Industry," Doctoral Dissertation, University of Wisconsin/Madison, May 1986 (L.R. Christensen, advisor), Appendix 4A, pp. 88-89. Note that the validity of the argument is not affected by the antiquity of the data.

23. The Uniform System of Accounts has been used by the FERC and its predecessors since 1938, as mandated by Congress.

24. See "Efficiency and Benchmarking Study of the NSW Distribution Businesses," IPART Research Paper No. 13, February 1999.

25. See Dykstra.

26. BNA (2006), 2. *Referenzbericht Anreizregulierung: Generelle sektorale Produktivitätsentwicklung im Rahmen der Anreizregulierung (2nd Reference BNA Report on Incentive Regulation: General sectoral productivity movements in the context of incentive regulation),* Bundesnetzagentur, Bonn, 26 January 2006.

27. See Dykstra, p. 1.

28. It is observed often enough that such formulation assumptions might not be appropriate in the case of a recently privatized company, with poorly main-

tained infrastructure, whose costs might be expected to fall faster than the "industry." That would be using the term "industry" too widely, however. It would not be practical to expect productivity growth for a newly privatized company to match that exhibited by a mature, investor-owned industry.

29. Note that Z^* can be positive or negative.

30. Recall that input price growth differs from output price growth by the growth in TFP. Only if national productivity growth were zero could a national output price index be a good measure of national input price growth.