Available online at www.sciencedirect.com
SCIENCE DIRECT
Energy Economics 26 (2004) 825–843
Energy Economics
www.elsevier.com/locate/eneco

Strategic behaviour under regulatory benchmarking

Tooraj Jamasb[a,*], Paul Nillesen[b,1], Michael Pollitt[c]

[a] *Department of Applied Economics, University of Cambridge, Sidgwick Avenue, Austin Robinson Building, Cambridge CB3 9DE, UK*
[b] *NUON NV, Netherlands*
[c] *Judge Institute of Management, University of Cambridge, UK*

Available online 17 June 2004

Abstract

In order to improve the efficiency of electricity distribution networks, some regulators have adopted incentive regulation schemes that rely on performance benchmarking. Although regulation benchmarking can influence the "regulation game," the subject has received limited attention. This paper discusses how strategic behaviour can result in inefficient behaviour by firms. We then use the Data Envelopment Analysis (DEA) method with US utility data to examine implications of illustrative cases of strategic behaviour reported by regulators. The results show that gaming can have significant effects on the measured performance and profitability of firms.

JEL classification: L94

Keywords: Gaming; Strategic behaviour; Regulation; Benchmarking; Electricity

1. Introduction

Since 1990, a liberalisation trend has transformed the structure, operating environment, and governance of the electricity sector in many countries around the world. An important aspect of this trend has been the establishment of regulatory agencies or, where a regulator already existed, a shift from rate-of-

* Corresponding author. Tel.: +44-1223-335271; fax: +44-1223-335299.
E-mail address: tooraj.jamasb@econ.cam.ac.uk (T. Jamasb).
[1] The views expressed are those of the author and do not represent the views of NUON NV.

0140-9883/$ - see front matter
doi:10.1016/j.eneco.2004.04.029

return (ROR) regulation of vertically integrated utilities, to regulation of diverse unbundled activities. In liberalised sectors, the potentially competitive generation and supply activities increasingly operate in a market-oriented environment. In addition, many transmission and distribution networks, generally viewed as natural monopolies, have undergone regulatory reform.

Opportunistic behaviour by regulated firms, including electric utilities, has been discussed extensively in the literature in the context of ROR regulation and asymmetric information between firm and regulator (see, e.g., Armstrong et al., 1994; Vickers and Yarrow, 1993). More recently, the notion of strategic behaviour by generating companies in the form of exercising market power in competitive wholesale electricity markets has attracted considerable interest. Market power in the generation market can arise from ownership concentrations, lack of access to and constraints in transmission networks, tight supply–demand conditions, and flawed trading and regulatory arrangements. This interest has arisen from the failure of some reforms to ensure effective competition, the recent electricity crisis in California, and, to some extent, from the collapse of the energy trading firm Enron (Borenstein et al., 2002; Joskow and Kahn, 2002).

In the post-reform era, some countries and jurisdictions have moved away from ROR regulation of transmission and distribution utilities and adopted incentive-based models. Some regulators, in particular those in Europe and Australia, have adopted benchmarking as a tool in the incentive regulation of network utilities (Jamasb and Pollitt, 2001). This development can affect the nature of the "regulation game" played between regulator and network utilities. However, this emerging aspect of regulatory gaming or strategic behaviour has received relatively little attention.

This paper focuses on strategic behaviour, or gaming, in the context of benchmarking in incentive regulation of distribution utilities. We refer to strategic behaviour or gaming as the type of behaviour that aims to increase profits without achieving real efficiency gains, i.e., they defy the incentive purpose of benchmarking, the regulatory objectives of efficient operation, and protection of public interest. It should be noted that "gaming" behaviour is not necessarily illegal and should be viewed within the regulatory context, as the optimisation process must remain within general accounting, fiscal, legal, and corporate governance statutes and policies.

In this study, we identify and examine the ways in which regulatory benchmarking can influence firm behaviour and analyse some possible implications. We then, utilising a data set of distribution activities of a sample of US electric utilities, illustrate strategic issues that a Public Utility Commission overseeing few electric utilities may encounter when using frontier-based benchmarking methods in incentive regulation. The purpose of the exercise is to examine the main issues involved and general lessons that are applicable to other regulatory settings.

The next section reviews the gaming aspects of regulatory benchmarking. Section 3 presents the data and methodology used in this study. Section 4 describes the main findings of a quantitative analysis of various gaming strategies on the outcome of regulatory benchmarking. Section 5 is a discussion of lessons and conclusions.

2. Gaming in incentive regulation benchmarking

2.1. Incentive regulation[2]

Asymmetric information between the regulator and the regulated firm is a key issue in the regulation of natural monopolies. Baron and Myerson (1982) and Laffont and Tirole (1986) address regulation of monopoly firms in the presence of asymmetric information in the form of unknown costs and unobservable effort to reduce costs.[3] A rather common criticism of the ROR regulation model is that it lacks incentives for efficiency improvements and encourages firms to engage in strategic behaviour. Averch and Johnson (1962) showed that ROR regulation encourages utilities to inflate their regulatory asset base through over-investment and socially inefficient resource allocation. The argument finds some parallels in the US power sector in the 1970s and 1980s where stranded costs of over-investment in generation capacity contributed to electricity price increases and, consequently, the calls for restructuring of the sector in the high-price states (Joskow, 1997).

Regulatory reform of network industries around the world has challenged the traditional ROR regulation, as regulators have adopted a variety of incentive-based models. These models aim to provide monopolies with the incentive to utilise their exclusive information on effort and costs to improve operating efficiency and investment decisions and to ensure that consumers benefit from the efficiency gains.[4] In the US, incentive-based regulation generally has taken the form of price cap or sliding scales (usually referred to as Performance-Based Regulation [PBR] or Rate-Making). The interest in incentive regulation is not due to new contributions from economic theory; rather, it reflects the need and desire for new practical approaches to regulation, even though these may not always be fully in line with theory (Crew and Kleindorfer, 1996, p. 215).

In this paper, we focus on price/revenue cap regulation model based on the RPI-X formula.[5] Price cap regulation de-couples profits from costs by setting maximum prices for the duration of a specified regulatory lag or rate period. The utility is then allowed to retain the profits in terms of the difference between the regulated price and its actual costs during the (typically 5-year) rate period. Price cap regulation was first implemented in post-privatisation regulation of British Telecom (Littlechild, 1983). The model has since been adopted in the regulation of other sectors in Britain and in many other countries.[6]

An important feature of incentive regulation is the use of benchmarking, which can be broadly defined as the comparison of a firm's actual performance against some predefined reference or benchmark performance. A perceived advantage of benchmarking has been that it reduces the information asymmetry problem that occurs in ROR

[2] This section draws heavily on Jamasb et al. (2003).

[3] See also Armstrong et al. (1994) for a review of these models.

[4] See Joskow and Schmalensee (1986) for a discussion of approaches to incentive regulation of electric utilities.

[5] For the purposes of this study, unless specified, we do not differentiate between a price and a revenue cap regulation based on the RPI-X formula.

[6] See Vickers and Yarrow (1993) for a description of the methodology and implementation in Britain.

regulation by reducing the regulator's reliance on the firm's own costs, but references the price to an external non-influencable benchmark.

2.2. The nature of regulation benchmarking game

A highly contentious issue in price and revenue cap regulation is how to determine efficiency improvements and translate these into tariff changes (X-factors). Regulators have adopted different benchmarking methods to arrive at X-factors, and it is in the implementation of these that the regulation game may be played. For the purposes of this study, we can distinguish between two types of benchmarking methods used in setting the X-factors: (i) frontier-based and (ii) non-frontier techniques.[7] The division also reflects the divide in benchmarking approaches used by, on the one hand, the European and Australian electricity regulators, and the PUCs in the United States on the other. The European regulators have generally adopted frontier-based methods as the basis on which to calculate the X-factors, while those PUCs that have adopted price caps have tended to use measures such as Total Factor Productivity (TFP) to calculate the efficiency requirements.

From a regulatory policy point of view, a major difference between the frontier and non-frontier benchmarking is that the former has a stronger focus on performance variations between firms. Frontier methods (such as Data Envelopment Analysis [DEA] and Corrected Ordinary Least Squares) appear suitable at initial stages of regulatory reform when a primary objective is to reduce the performance gap among the utilities through firm-specific efficiency requirements. Non-frontier methods may be used to mimic competition among firms with relatively similar cost levels, or when there is a lack of data and comparators for the use of frontier methods. There is also an important methodological difference between frontier and non-frontier approaches. In the frontier-based approaches, the efficiency scores are measured relative to an efficient frontier. This results in an interdependence between a firm's efficiency measure (score) and strategic behaviour involving frontier firms in the sample. In the index number approach to TFP at sector level, each firm's benchmark is the same and can only be marginally (if at all) affected by own or other firms' strategic behaviour.

In principle, the aim of benchmarking within incentive regulation is to exploit the efficiency improvement potential of the regulated firm. Regulators should recognise that their benchmarking exercise inevitably shapes the efforts and directs considerable resources of the firms towards the make-up and variables of these models. However, while benchmarking can measure "true" performance improvements, gaming can sometimes produce illusive or "virtual" efficiency improvements. Therefore, benchmarking models need to strike a balance between reflecting the main performance drivers of the business in question and reducing incentives for engaging in unproductive method or model-induced strategic behaviour.

This type of behaviour is rational from a firm's perspective. Optimising the regulatory process and exploiting the information advantage will maximise profits for shareholders. In cases where customers are, directly or indirectly, shareholders (e.g., cooperatives or

[7] The review of the methods in this section is based on Jamasb and Pollitt (2001) and Pollitt (1995). See also Coelli et al. (1998) and DTe (1999).

mutuals in the strict sense, or municipal owned), the firm's excess profits might still benefit the local consumer. However, where customers have no relation with the capital of the firm, such regulatory strategies are likely to lead to welfare losses.

Regulated firms may attempt to influence the use of regulation benchmarking at the adoption stage. Although these efforts may not be considered as gaming, utilities may attempt to influence (i) the use of benchmarking in incentive regulation, (ii) the choice of method, model, and variables (and their weighting), (iii) the definition of variables adopted during the consultation process, and (iv) the translation of efficiency scores into *X*-factors. At a later stage, firms may use gaming strategies to benefit from the regulator's adopted benchmarking model.

Some regulation games are associated with the periodic aspect of ROR and incentive-based regulatory reviews through timing of specific types of actions. Dynamic aspects of strategic behaviour of the firm associated with regulatory lag are known to regulators and have been addressed by some authors (see Baumol and Klevorick, 1970; Sappington, 1980). Di Tella and Dyck (2002), in a study of the Chilean electricity distribution utilities under price cap regulation, report evidence of cyclical cost reductions that coincide with the initial years of rate periods and the reverse prior to the next rate review.

Gaming behaviour is not only limited to private firms. Publicly owned firms can also be motivated to pursue monetary or other performance measures. Several countries noted for the use of benchmarking, including the Netherlands, Norway, and Australia, have significant municipal or state ownership. Courty and Marschke (2002), in a study of job training agencies, show that public organisations can engage in gaming by timing their performance reports in order to benefit from awards. They show that performance incentives can come at a cost by having a negative effect on efficiency.

Broadly, it is possible to differentiate between two types of strategic behaviour. The first is behaviour that may not have a material effect on the efficient operation of the firm and is intended to present the performance of the firm in a more favourable light. For example, a firm may shift costs from operating to capital costs or influence the choice of output variables in order to affect measured relative performance. The main undesirable outcome of such virtual efficiencies is that they result in welfare transfer from customers, or even other firms, to the gaming firm through lower efficiency targets than the true underlying efficiency would suggest.

The second type of gaming is in the form of behaviour that distorts the efficient operation and investment decisions of the firm. For example, the firm might increase its cost base or delay efficiency improvements in periods leading to a new rate case. This type of gaming results in socially inefficient resource allocation and dead-weight loss. An important concern with both of these gaming categories in frontier-based approaches is that, due to the interdependency between the efficiency scores, a firm's gaming can also affect the measured performance of other firms.

In cost-based DEA models, the regulator may use controllable operating expenditure (OPEX) as the input variable and treat capital expenditures (CAPEX) outside of the benchmarking exercise such as in the UK or alternatively use total expenditures (TOTEX) as in The Netherlands. This has implications for the possible strategic behaviour by firms. A firm may appear more efficient by reducing costs, as well as by appearing larger in terms of higher output variables. For example, if the benchmarking model uses OPEX as input and network length or transformer capacity as output and is given approval for expansion plans for increasing separately treated CAPEX, it can earn a return on its capital expenditures and, at the same time, increase the output variables and hence its relative efficiency.

Alternatively, the regulator may use OPEX and CAPEX as two separate variables. This will allow the possible trade-offs between the two types of costs to be reflected in the model. It has been suggested that where OPEX is used as the only input variable, the chosen outputs should be independent of CAPEX or adjusted to reflect their relative share of total costs (Coelli, 2000). In general, a complicating factor in benchmarking is a question of which model specification best represents the activity of electric distribution utilities. Jamasb and Pollitt (2001) show that efficiency studies of distribution utilities have used a variety of variables and model specifications, and the issue is not yet satisfactorily settled.[8]

The choice of benchmarking model could also serve specific regulatory objectives. For example, the choice of constant or variable returns to scale models can affect the long-term structure of the sector. The Dutch regulator has used a constant returns-to-scale DEA model and assumes that firms can freely adjust their scale of operations through mergers and acquisitions. Other countries value the maintenance of the number of comparators and use variable returns to scale measures (e.g., the UK). The UK regulator has estimated the cost of information loss due to mergers among electricity distribution utilities at £32 million and applies a corresponding reduction in regulated revenue of the merged firms over a 5-year period (OFGEM, 2002).

2.3. A survey of electricity regulators

As discussed in the previous section, a number of national and state-level electricity regulators have now adopted regulatory benchmarking using frontier methods. As part of the background research for this paper, we conducted a survey of eight of these electricity regulators in order to understand their experience with strategic behaviour associated with regulatory benchmarking. The full survey results appear in Jamasb et al. (2003); we briefly review them here.

The survey returns reveal that regulators have experienced three major types of gaming strategies. The first category includes possible strategic behaviour that is associated with cost issues. All eight of the surveyed regulators had experienced potential gaming in the form of shifting of costs and assets across sectors (e.g., electricity vs. gas or water) and within the electricity sector (e.g., generation, transmission, distribution, and supply), costing rules, definitions, and rate of return by firms. The second category of reported issues (experienced again by all eight regulators) involved possible gaming of the methodology used by the regulators such as influencing the use of benchmarking models to be used, choice of output and input variables (and their weighting), and information disclosure. The third category (reported by two of the regulators) is concerned with utility mergers, an issue this is increasingly faced by some regulators. We follow a rather similar classification in the analysis of effects of three selected cases of possible strategic behaviour (in Sections 4.2–4.4).

[8] In extremis, certain variables may be used as inputs in one regulatory model, whereas in other regimes, they are used as outputs. For a Dutch discussion, see Nillesen and Telling (2001).

It should be noted that determining whether certain behaviour by firms constitutes gaming is to some degree a subjective matter and therefore requires judgement. In other words, the perceived motives are often observed indirectly through their effect on the regulatory objectives or outcomes. Further, because of the complex interplay of firms and issues, it is virtually impossible to isolate and predict the outcome of a particular strategy of an individual firm, making it difficult to separate cause and effect. All gaming opportunities must be conducted within the prevailing legal, accounting, fiscal, and corporate governance regulations, so gaming should be seen as regulatory model optimisation rather than fraudulent or deceptive behaviour.

3. Data and methodology

3.1. Data

In order to illustrate numerical examples of the possible effects of strategic behaviour in regulatory benchmarking, we utilise a data set comprising electricity distribution business of 28 utilities operating in the northeast of the United States. The data used is based on annual company returns to the Federal Energy Regulation Commission (FERC) and Platts (2002) for the financial year 2000. We focus on a subset of five firms in our sample and examine the effects of gaming on these. These focus firms could be viewed as firms operating under the jurisdiction of a regulatory commission in a federal state being benchmarked against a national sample.

Alternatively, the focus firms may be regarded as utilities operating under a national regulator being benchmarked against an international sample.[9] Individual utilities are not identified here, as we wish to address the issues involved at a general level. Table 1 shows the summary statistics for the entire sample and for the focus firms. As shown in the table, the focus firms are relatively far from the extreme ends of the larger sample.

3.2. Data envelopment analysis

For the purposes of illustration, we assume that our example regulator makes use of a frontier benchmarking technique, data envelopment analysis (DEA) in the setting of firm-specific *X*-factors. DEA has been a popular benchmarking method with electricity regulators (see Jamasb and Pollitt, 2001). DEA identifies an efficient frontier made up of the best practice firms and uses this to measure the relative efficiency scores of the less efficient firms. Norway uses the DEA in setting revenue caps for regional electricity transmission and distribution utilities. An advantage of the method is that it does not require specification of a production or cost function. It allows calculation of allocative and technical efficiencies that can be decomposed into scale, congestion, and pure technical efficiencies (Färe et al., 1985).

DEA is a nonparametric method and uses piecewise linear programming to calculate (rather than estimate) the efficient or best practice frontier of a sample (Farrell, 1957; Färe

[9] See Jamasb and Pollitt (2003) for an example and discussion of international regulation benchmarking.

Table 1
Summary statistics for the sample and focus firms

	Distribution OPEX (× US$1000)*	Electricity delivered (MWh)	Number of customers	Network length (miles)	Number of meters	Maximum demand (MW)
Focus firms						
Firm 1	88,033	14,144,052	582,339	16,390	252,622	2673
Firm 2	219,238	21,261,331	1,000,526	23,391	1,052,369	4961
Firm 3	43,608	14,607,563	491,142	14,900	524,605	2342
Firm 4	20,057	7,933,735	134,554	6121	174,067	926
Firm 5	94,822	21,714,983	680,405	21,735	789,637	3311
Sample						
Mean	110,873	13,505,020	553,329	13,095	598,271	2324
Minimum	2885	124,425	26,672	120	27,840	101
Maximum	478,345	41,834,169	3,074,592	41,000	3,301,863	9379

*The operating expenditures are calculated from reported data and adjusted to allow for an allocation of common administration costs.

et al., 1985). The decision-making units (DMUs) or firms that make up the frontier envelop the less efficient firms. The efficiency of the firms is calculated in terms of scores on a scale of 0 to 1, with the frontier firms receiving a score of 1.

DEA models can be output or input oriented and can be specified as constant returns to scale (CRS) or variable returns to scale (VRS). Output-oriented models maximise output for a given amount of input. Conversely, input-oriented models minimise input factors required for a given level of output. An input-oriented specification is generally regarded as the appropriate form for electricity distribution utilities, as demand for their services is a derived demand that is beyond the control of utilities and that has to be met.

The linear program calculating the efficiency score of the ith firm in a sample of N firms in CRS models takes the form specified in Eq. (1) where θ is a scalar (equal to the efficiency score) and λ represents an $N \times 1$ vector of constants. Assuming that the firms use K inputs and M outputs, X and Y represent $K \times N$ input and $M \times N$ output matrices, respectively. The input and output column vectors for the ith firm are represented by x_i and y_i, respectively. The equation is solved once for each firm. In VRS models, a convexity constraint $\Sigma\lambda = 1$ is added. This additional constraint ensures that the firm is compared against other firms with similar size.

$$\begin{aligned} &\min_{\theta,\lambda} \theta, \\ &\text{s.t.} \\ &-y_i + Y\lambda \geq 0, \\ &\theta x_i - X\lambda \geq 0, \\ &\lambda \geq 0 \end{aligned} \tag{1}$$

In Eq. (1), firm i is compared to a linear combination of sample firms which produce at least as much of each output as it does with the minimum possible amount of inputs. Fig. 1

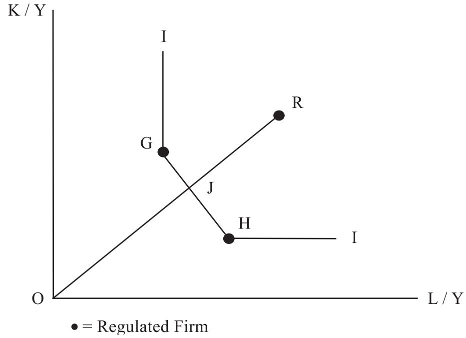

Fig. 1. Data envelopment analysis.

illustrates the main features of an input-oriented model with constant returns to scale. The figure shows three firms (G, H, R) that use two inputs (capital K, labour L) for a given output Y. The vertical and horizontal axis represent the capital and labour costs per unit of output, respectively.

Firms G and H produce the given output with lower inputs and form the efficient frontier that envelops the less efficient firm R. The technical efficiency of firm R relative to the frontier can be calculated from OJ/OR ratio. Technical efficiency measures the ability of a firm to minimise inputs to produce a given level of output.

An important step in DEA is the choice of appropriate input and output variables. The variables should, to the extent possible, reflect the main aspects of resource use in the activity concerned. DEA can also control for the effect of environmental variables that are beyond the control of the management of firms but affect their performance. In addition, the basic DEA model illustrated above does not impose weights on model input and output variables. However, the model can be extended to incorporate value judgements in the form of relative weight restrictions imposed on model inputs or outputs. This can be achieved by including additional constraints to the model. The aim is to control for the influence of values of individual input and outputs on the efficiency scores (Thanassoulis, 2001).

An advantage of DEA is that inefficient firms are compared to actual firms rather than to a statistical measure. In addition, DEA does not require specification of a cost or production function. However, efficiency scores tend to be sensitive to the choice of input and output variables. Furthermore, the results (scores) are sensitive to measurement errors in the frontier firms as these comprise the best practice frontier. Also, the method does not allow for stochastic factors and measurement errors. Finally, as more variables are included in the models, the number of firms on the frontier increases; therefore, it is important to examine the sensitivity of the efficiency scores and rank order of the firms to model specification.

3.3. Preferred models

In order to examine the possible effects of strategic behaviour on the outcome of regulation benchmarking, we use relatively familiar DEA model specifications. An initial model serves as the reference or base model against which we compare the outcomes of strategic behaviour. A financial model is also used to calculate the benefits and losses from changes in relative efficiency.

3.3.1. DEA models

Our preferred model is input oriented and assumes constant returns to scale (CRS), so that the measured relative efficiency of firms is not affected by their size. This is consistent with the DEA models adopted by the Dutch and Norwegian regulators. Empirical studies in Norway, Canada, New Zealand, and Switzerland find evidence of the presence of economies of scale in electric distribution utilities. In some studies, the minimum efficient firm size in terms of number of customers is estimated to be around 20,000–30,000.[10] In addition, Allas and Leslie (2001) report that about 85% of costs vary with the number of customers and the units of energy delivered.

The preferred model uses a single cost input reflecting the OPEX of the distribution business of the utilities. The output variables in our preferred model are (i) units of electricity delivered, (ii) number of customers, and (iii) length of network. The literature on relative efficiency analysis and benchmarking does not reveal a universally agreed set of input and output variables for modelling of electricity distribution utilities. However, as reported in Jamasb and Pollitt (2001), the input and output variables in our simple model are among the most widely used in studies of relative performance.

3.3.2. Financial model

The financial model calculates the efficient level of costs for individual firms as the product of the efficiency scores and OPEX in the reference year. The model assumes that the efficient cost levels are achieved by the end of a 5-year regulatory rate period and that the relative efficiency score accurately reflects the possible cost savings without incorporating measurement error or other stochastic components. This is achieved through annual efficiency improvement requirements or X-factors calculated for individual firms. Fig. 2 illustrates a gliding path reflecting combinations of X-factors and reference prices. In this example, a firm's OPEX are benchmarked, while it is allowed to recover its depreciation costs and earn a weighted average cost of capital (WACC) on its regulatory asset base (RAB). In our examples we have used a WACC of 6.8%. Line AC shows the path bringing the allowed revenues of the firm to the efficient frontier level, inclusive of an anticipated frontier shift BC, during the course of a 5-year rate period.

However, in reality, the ability of highly inefficient firms to achieve cost savings during a given period may be limited. Recognising this practical limitation, regulators in the Netherlands and Norway have introduced limits on maximum efficiency requirements imposed on least efficient firms. The assigned X-factors in this model are therefore capped at 8% per year (as in the Netherlands). In addition, the model assumes that the efficient cost levels can achieve a further 1% efficiency gain (frontier shift) per year over a 5-year rate period.

[10] See Filippini (1996, 1997, 1998) and Filippini and Wild (1998) for studies of Switzerland; Salvanes and Tjotta (1994, 1998) for Norway; Giles and Wyatt (1993) for New Zealand; and Yatchew (2000) for Canada.

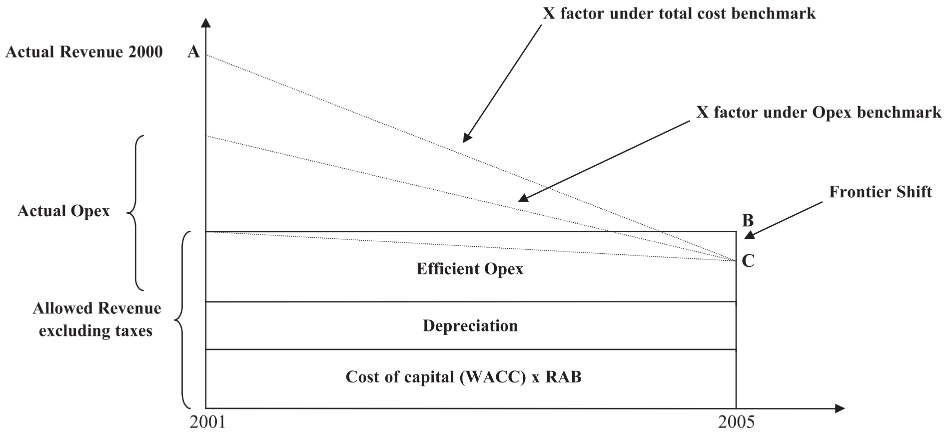

Fig. 2. The gliding path of allowed revenue during a rate period.

The total efficiency requirements or effective X-factors are then used to calculate the allowed revenues to cover the firms' operating expenditures. Ordinarily, regulatory models calculate X-factors for total allowed distribution revenues regardless of whether operating or total expenditures (operating plus capital expenditure) are benchmarked. However, for the purposes of simplicity and transparency, we focus solely on the gaming OPEX and its effect on efficiency score and X-factor.

4. Results

As highlighted previously, there are various ways in which strategic behaviour can affect the outcome of a regulation benchmarking. In this section, we report benchmarking results using a base model, which we assume to be the regulator's model of choice in the absence of gaming behaviour. We then examine the results of three selected cases (based on our survey of electricity regulators) of deviation from the base model that can arise from strategic gaming.

4.1. Base case—no gaming

Table 2 shows the calculated efficiency scores for the distribution business of 28 electric utilities in our sample (the base model was described in Section 5). Utilities F1–F5 are the focus firms assumed to be operating under the jurisdiction of a single regulator and being benchmarked within a sample of firms. As shown in the sample, the range of efficiency scores for the sample is rather wide (26–100%). Four firms, two of which are among our focus firms (F3 and F4), have an efficiency of 100% and constitute the efficient frontier. Utilities F1 and F2 of the focus firms score relatively low, while firm F5 is the most efficient non-frontier firm in our focus group.

Table 3 summarises the results of the base model. As shown in the table, the implied X-factor for firms F1 and F2 exceeds the maximum 8% and is therefore capped at that

Table 2
Efficiency scores for the sample—base case

Number	Efficiency score (%)
F1	57.6
F2	38.1
F3	100.0
F4	100.0
F5	67.6
F6	41.6
F7	29.0
F8	30.4
F9	78.9
F10	32.1
F11	43.2
F12	43.7
F13	61.6
F14	66.8
F15	100.0
F16	72.8
F17	34.9
F18	22.6
F19	58.5
F20	49.3
F21	41.8
F22	34.3
F23	50.8
F24	60.0
F25	76.1
F26	57.3
F27	100.0
F28	43.6

level. The final *X*-factor is the effective (or total) rate of cost reduction assigned to the firms and includes a 1% annual frontier shift in efficient cost level. It is interesting to note that under a capping regime, a lower efficiency score can translate into a lower final *X*-factor.

Table 3
Summary results—base case

Base case	Firm 1	Firm 2	Firm 3	Firm 4	Firm 5
Efficiency score (%)	57.6	38.1	100.0	100.0	67.6
X-factor (%)	10.4	17.5	0.0	0.0	7.5
Implied *X*-factor (%)	8.0	8.0	0.0	0.0	7.5
Final *X*-factor (%)	8.8	8.5	1.0	1.0	8.5
OPEX reference year (× US$1000)	88,033	219,238	43,608	20,057	94,822
Accumulated required cost savings rate period (× US$1000)	102,537	249,027	6455	2969	106,575
Accumulated allowed OPEX rate period (× US$1000)	337,629	847,161	211,587	97,316	367,535

The two frontier firms in our focus group are the smallest firms in terms of OPEX in the reference year, while the largest firm appears as the least efficient firm among these. The table also shows the allowed OPEX for the firms during the rate period and the required cost reductions that the effective *X*-factors represent in relation to the OPEX in the reference year (undiscounted).

4.2. Gaming OPEX

As discussed previously, one of the perverse incentives associated with price cap regulation is that firms may attempt to inflate their distribution cost base before a price review. As shown here, relatively small changes in *X*-factors that do not appear significant in relation to the original *X*-factor can result in considerable revenue implications.

Table 4 shows the effect of a 1%, 5%, and 10% increase in the OPEX of firm F5 in the reference year 2000 on its allowed revenue for the rate period 2001–2005. The table shows that as the firm's cost base increases, its efficiency score declines. This results in higher *X*-factors and consequently higher cost-saving requirements. At the same time, the firm enjoys a higher cost base that gives it a higher level of allowed revenues. The table shows the net increase in allowed cost recovery or revenue after controlling for the effect of higher efficiency requirements through higher *X*-factor. For example, a 5% increase in the cost base prior to the rate review results in a reduction of the efficiency score of 6 percentage points, yet also results in a 3.6% net increase in allowed revenues corresponding to $19.6 per customer for the rate period.

Table 5 shows the effect of a 10% cost inflation by each of the other four firms for the other four firms (keeping each of the other firms' costs unchanged). As shown in the table,

Table 4
Summary results—gaming by increasing OPEX

Firm 5	Base case	Base OPEX plus 1%	Base OPEX plus 5%	Base OPEX plus 10%
Efficiency score (%)	67.6	66.9	64.3	61.4
X-factor (%)	7.5	7.7	8.4	9.3
Implied *X*-factor (%)	7.5	7.7	8.0	8.0
Final *X*-factor (%)	8.5	8.7	8.9	8.9
OPEX reference year (× US$1000)	94,822	95,770	99,563	104,304
Accumulated required cost savings (× US$1000)	106,575	109,697	116,954	122,072
Δ accumulated required cost savings (× US$1000)	–	3122	10,379	15,497
Accumulated allowed OPEX in rate period (× US$1000)	367,535	369,154	380,861	399,449
Δ accumulated allowed OPEX in rate period (× US$1000)	–	1,620	13,327	31,914
As % of accumulated allowed revenue (in base case) (%)	–	0.4	3.6	8.7
$ Revenue increase (+)/decrease (−) per customer	–	2.38	19.59	46.90

Table 5
Summary results—gaming by increasing OPEX by 10%

	Firm 1 base OPEX plus 10%	Firm 2 base OPEX plus 10%	Firm 3 base OPEX plus 10%	Firm 4 base OPEX plus 10%
Efficiency score (%)	52.4	34.7	100.0	100.0
Δ score from base case (%)	− 5.2	− 3.4	0	0
X-factor (%)	12.1	19.1	0.0	0.0
Implied *X*-factor (%)	8.0	8.0	0.0	0.0
Final *X*-factor (%)	8.7	8.5	1.0	1.0
OPEX reference year (× US$1000)	96,837	241,161	47,969	22,063
Accumulated required cost savings (× US$1000)	112,040	272,694	7100	3266
Δ accumulated required cost savings (× US$1000)	9503	23,668	645	297
Accumulated allowed revenue in rate period (× US$1000)	372,143	933,112	232,746	107,048
Δ accumulated allowed OPEX in rate period (× US$1000)	34,514	85,951	21,159	9732
As % of accumulated allowed OPEX (in base case) (%)	10.2	10.1	10.0	10.0
$ Revenue increase (+)/decrease (−) per customer	59.27	85.91	43.08	72.33

the frontier firms F3 and F4 retain the full increase in their cost base. The least efficient firms F1 and F2, despite receiving lower efficiency scores relative to the base case, benefit fully from higher cost base due to their capped *X*-factors. They also achieve an additional small gain as the 1% frontier shift is applied to a lower efficient cost base.

4.3. Influencing output weights

The selection of appropriate output variables for use in benchmarking models can be a source of disagreement between regulator and firm. In principle, the main issue is the extent to which the selected variables are true cost drivers and how accurately they portray the production function. The underlying concern is how the choice and use of variables can affect the relative efficiency measure and, consequently, the firms' revenues. We modify our base model by assigning a set of weights to the output variables.

As mentioned in Section 4, DEA can be used with weight restrictions applied to inputs and outputs. In the following example, we examine the effect of output weights similar to those used by the UK regulator OFGEM, namely, (i) number of customers 50%, (ii) units of electricity delivered 25%, and (iii) length of distribution network 25%. The weights on the outputs are introduced by including additional constraints to our basic DEA model (see Thanassoulis, 2001).

Table 6 shows the summary results of applying the weights to the output variables in the base model. As shown in the table, after introducing the output weights, firm F3 is no longer on the efficient frontier and its allowed revenues decrease by 7.5% corresponding to a $32.4 revenue reduction per customer. It is noteworthy that although firms F1 and F2

Table 6
Summary results—influencing output with fixed weights

Outputs with fixed weights	Firm 1	Firm 2	Firm 3	Firm 4	Firm 5
Efficiency score (%)	42.5	26	87.4	100.0	59.5
Δ score from base case (%)	− 15.1	− 12.1	− 12.6	0	− 8.1
X-factor (%)	15.8	23.6	2.7	0.0	9.9
Implied *X*-factor (%)	8.0	8.0	2.7	0.0	8.0
Final *X*-factor (%)	8.6	8.4	3.6	1.0	8.8
OPEX in reference year (× US$1000)	88,033	219,238	43,608	20,057	94,822
Accumulated required cost savings (× US$1000)	100,568	245,074	22,382	2969	110,708
Δ accumulated required cost saving (× US$1000)	− 1969	− 3952	15,928	0	4133
Accumulated allowed revenue in rate period (× US$1000)	339,598	851,114	195,659	97,316	363,402
Δ accumulated allowed revenue in rate period (× US$1000)	1969	3952	− 15,928	0	− 4133
As % of accumulated allowed OPEX (in base case) (%)	0.6	0.5	− 7.5	0.0	− 1.1
$ Revenue increase (+)/decrease (−) per customer	3.38	3.95	− 32.43	0.0	− 6.07

have lower efficiency scores than in the base case, they achieve a relatively small net gain from the weights. The observed gain is due to the fact that in the revised model the 1% frontier shift is applied to a lower efficient cost base while the firms' *X*-factor is, despite a nominal increase, still capped at 8%. However, a lower efficiency score for firm F5 means that it faces an increase in effective *X*-factor up to the cap limit that results in a net loss for the firm.

The results shown here indicate that while there is some potential for moderate gains by the two least efficient firms, the negative effect on more efficient firms outweighs the gain. The results from this example indicate that conflicting interest among the firms may reduce the likelihood of influencing the base model in this direction. It should, however, be pointed out that the magnitude of potential benefits and losses, and the number of gainers and losers, is highly dependent on the composition of the companies comprising the benchmarking sample.

4.4. Changing the firm through mergers and acquisitions

The changing structure of the electricity industry has prompted many utilities to achieve efficiency improvements through mergers and acquisitions. From a strategic perspective, mergers can also help utilities to reposition themselves in the market by changing their scale of operations and reconfiguring their resources. However, mergers and acquisitions involve two sources of concern for regulators that use benchmarking: (i) transactions intended to influence the relative position of the firm without achieving real efficiency gains and (ii) the shrinking number of firms and reduction in information on which regulators base their analysis (Nillesen et al., 2001). The first type of merger

Table 7
Summary results—merger effects

Effects of merger between firms F3 and F5	Firm 1	Firm 2	Firm 4	Firms 3 and 5
Efficiency score (%)	67.3	42.7	100.0	91.5
Δ score from base case (%)	9.7	4.6	0	18.2
X-factor (%)	7.6	15.6	0.0	1.8
Implied *X*-factor (%)	7.6	8.0	0.0	1.8
Final *X*-factor (%)	8.5	8.6	1.0	2.7
OPEX in reference year (× US$1000)	88,033	219,238	20,057	138,430
Accumulated required cost savings (× US$1000)	99,621	250,510	2969	54,506
Δ accumulated required cost saving (× US$1000)	− 2916	1483	0	− 58,523
Accumulated allowed revenue in rate period (× US$1000)	340,545	845,678	97,316	637,645
Δ accumulated allowed revenue in rate period (× US$1000)	2,916	− 1,483	0	58,523
As % of accumulated allowed OPEX (in base case) (%)	0.9	− 0.2	0.0	10.1
$ Revenue increase (+)/decrease (−) per customer	5.01	− 1.48	0.00	49.95

may be regarded as a special form of collusion to game the regulator's incentive scheme.[11]

In this section, we examine a case of "virtual" efficiency improvement achieved by the merger of two firms. Table 7 shows the results when a frontier firm (F3) and a relatively efficient firm (F5) merge to form a new entity.

Firms F1 and F2 exhibit higher efficiency scores relative to the base case due to the reduction in the number of frontier firms. A higher score for F1 means that the firm's *X*-factor falls below the cap threshold and therefore benefits from the new higher efficiency score. However, for firm F2, the higher efficiency score means that the 1% annual frontier shift is applied to a higher cost base, while the firm's *X*-factor remains capped despite the higher efficiency scores.

A merger between firms F3 and F5 can result in 10.1% increase in allowed revenues corresponding to $50 per customer during the rate period or $59 million in total. Although a range of factors can influence a firm's decision to merge or acquire, simple comparative efficiency analysis can reveal the side benefits or losses associated with the decisions. This type of analysis could indicate the premium a particular firm might be willing to pay compared to that of other firms.

5. Discussion and conclusions

An increasing number of regulators have used benchmarking in periodic price controls as part of the incentive-based regulation of natural monopolies. The primary reason for the

[11] Collusion can also take the form of collaboration on the pace of cost saving effort among the firms but this is beyond the scope of this study (see e.g. CPB, 2000).

use of benchmarking has been that yardstick regulation encourages efficiency and reduces reliance on the firms' own information. However, as discussed in this paper, the use of benchmarking can lead firms to pursue virtual rather than true performance improvements by gaming the regulator's benchmarking in a number of ways that are contrary to the intentions of the scheme.

We showed how strategic behaviour in the context of benchmarking may lead to (i) foregone efficiency improvements or dead-weight losses, (ii) welfare transfers from customers to firms, and (iii) welfare transfers among firms. We also used numerical examples to illustrate selected aspects of strategic gaming associated with regulatory benchmarking and their effects. We show that the net effect of gaming can depend on the method of translating efficiency scores into X-factors and the caps applied to them (i.e., the extent to which the efficiency gap among the firms can or is to be closed in a given rate period). We also showed the interrelationship between gaming by one firm and its effect on the X-factors for other firms.

The following lessons can be drawn from our review of issues and examples:

- The allocation of costs and assets when distribution is unbundled from other utility activities and the reliability of this information base for subsequent price controls are both important. Regulators need to pay particular attention to increasing the reliability of information through audits, technical studies, and comparison of cost patterns in review vs. non-review periods.
- Regulators need to conduct sensitivity analyses of their chosen benchmarking approach and data sets in order to identify the most influential variables and to assess the effects of measurement errors and likely gaming.
- An important strength of DEA is the ability to accommodate multiple inputs and outputs. Using models with a single cost input variable might increase the sensitivity of results to changes in costs.
- Mergers are increasingly a source of concern for regulators and utilities. Both can use benchmarking analyses to determine the effects of possible and actual mergers on the firms in the sector and their implied X-factors. Such analysis can also help regulators to design their policies towards separating virtual from actual efficiency gains in mergers.
- Regulators need to recognise the shortcomings of their chosen benchmarking methods and to apply discretion and judgment in the use of results. For example, in order to reduce reliance on a limited number of variables, regulators can use competing models and average their results. In some instances, it may be preferable to simplify the process by placing the firms in a few categories with similar X-factors.
- Finally, from a theoretical and methodological point of view, regulatory benchmarking leaves considerable scope for improvement. Regulatory benchmarking therefore owes much of its legitimacy to the wider regulatory framework and the implementation process. Transparency of the benchmarking exercise and decision process together with the public availability of underlying data and, combined with consultations and hearings, can provide third party scrutiny and thus increase acceptability and reduce gaming.

Regulatory benchmarking does not eliminate the issue of asymmetric information on firms' costs and efficiency improvement effort as known under rate of return regulation.

Rather, it adds new dimensions to this issue and the ways in which firms can behave strategically. Countering strategic behaviour can partly be overcome by increasing data accuracy and improving data collection procedures. The information requirement for reliable regulatory benchmarking therefore appears to be higher than initially expected. The continued efforts made by regulators using benchmarking to improve data quality are testament to this fact. At the same time, regulated utilities need to conduct their own benchmarking analysis in order to

- examine the effect of the regulator's choice of method, variables, *X*-factors,
- analyse the effects of possible gaming by other firms and available future partners for mergers and acquisitions, and
- evaluate benefits and losses of mergers involving own firm or competitors and to convey their findings to regulators.

Acknowledgements

The authors acknowledge the financial support of CMI Project IR-45 and thank Liz Hooper for her help with this paper. The usual disclaimer applies.

References

Allas, T., Leslie, K., 2001. Sizing power. The McKinsey Quarterly 1, 138–145.

Armstrong, M., Cowan, S., Vickers, J., 1994. Regulatory Reform: Economic Analysis and British Experience. The MIT Press, Cambridge, MA.

Averch, H., Johnson, L.L., 1962. Behavior of the firm under regulatory constraint. American Economic Review 52, 1052–1069 (December).

Baron, D.P., Myerson, R.B., 1982. Regulating a monopolist with unknown costs. Econometrica 50, 911–930.

Baumol, W.J., Klevorick, A.K., 1970. Input choices and rate of return regulation. Bell Journal of Economics 1, 162–190.

Borenstein, S., Bushnell, J., Wolak, F., 2002. Measuring market inefficiencies in California's restructured wholesale electricity market. American Economic Review 92 (5), 1376–1405 (December).

Coelli, T., 2000. Some scattered thoughts on performance measurement for regulation of a natural monopoly network industry, mimeo. Centre for Efficiency and Productivity Analysis, School of Economic Studies, University of New England.

Coelli, T., Rap, D.S.P., Battese, G.E., 1998. An Introduction to Efficiency and Productivity Analysis. Kluwer Academic Publishing, London.

Courty, P., Marschke, G., 2002. An empirical investigation of gaming responses to explicit performance incentives. Center for Economic Policy Research (CEPR). Discussion Paper 3164. January.

CPB, 2000. Yardstick competition: Theory, design, and practice, Netherlands Bureau for Economic Policy Analysis. Working Paper, No. 133, The Hague, (December).

Crew, M.A., Kleindorfer, P.R., 1996. Incentive regulation in the United Kingdom and the United States: some lessons. Journal of Regulatory Economics 9, 211–225.

Di Tella, R., Dyck, A., 2002. Cost reductions, cost padding and stock market prices: the Chilean experience with price cap regulation. Working Paper 03-050, Harvard Business School.

DTe, 1999. Price Cap Regulation in the Electricity Sector. Netherlands Electricity Regulatory Service (DTe), Hague.

Färe, R., Grabowski, R., Grosskopf, S., 1985. The Measurement of Efficiency of Production. Kluwer Academic Publishing, Boston.

Farrell, M.J., 1957. The measurement of productive efficiency. Journal of Royal Statistical Society, Series A 120, 253–281.
Filippini, M., 1996. Economies of scale and utilization in the Swiss electric power distribution industry. Applied Economics 28, 543–550.
Filippini, M., 1997. Elements of the Swiss Market for Electricity Physica-Verlag, Berlin.
Filippini, M., 1998. Are municipal electricity distribution utilities natural monopolies? Annals of Public and Cooperative Economics 2, 157–174.
Filippini, M., Wild, J., 1998. The estimation of an average cost frontier to calculate benchmark tariffs for electricity distribution. Working Paper No. 9803. Socioeconomic Institute, University of Zurich. (September).
Giles, D., Wyatt, N.S., 1993. Economies of scale in New Zealand electricity distribution industry. In: Phillips, P.C.B. (Ed.), Models, Methods and Applications of Econometrics. Blackwell, Oxford, pp. 370–382.
Jamasb, T., Pollitt, M., 2001. Benchmarking and regulation: international electricity experience. Utilities Policy 9 (3), 107–130.
Jamasb, T., Pollitt, M., 2003. International benchmarking and regulation: an application to European electricity distribution utilities. Energy Policy 31 (15), 1609–1622 (December).
Jamasb, T., Nillesen, P., Pollitt, M., 2003. Gaming the regulator: a survey. The Electricity Journal 16 (10), 68–80 (December).
Joskow, P.L., 1997. Restructuring, competition and regulatory reform in the US electricity sector. Journal of Economic Perspectives 11 (3), 119–138 (Summer).
Joskow, P.L., Kahn, E., 2002. A quantitative analysis of pricing behaviour in California's wholesale electricity market during summer 2000. The Energy Journal 23 (4), 1–36.
Joskow, P.L., Schmalensee, R., 1986. Incentive regulation for electric utilities. Yale Journal on Regulation 4 (1), 1–49.
Laffont, J.-J., Tirole, J., 1986. Using cost information to regulate firms. Journal of Political Economy 94, 614–641.
Littlechild, S., 1983. Regulation of British Telecommunications' Profitability. Department of Industry, London.
Nillesen, P.H.L., Telling, J.G., 2001. Benchmarking distribution companies. Energy and Power Risk Management, March, 10–13.
Nillesen, P.H.L., Pollitt, M.G., Keats, K., 2001. Strategic benchmarking: identifying and quantifying the gains from M&A. Energy and Power Risk Management, June, 42–45.
OFGEM, 2002. Mergers in the electricity distribution sector. Policy Statement. Office of Gas and Electricity Markets (OFGEM). London, (May).
Platts, 2002. Platts directory of electric power producers and distributors. 110th Edition of the Electrical World Directory. The McGraw-Hill Companies, New York.
Pollitt, M., 1995. Ownership and Performance in Electric Utilities. Oxford Univ. Press, Oxford.
Salvanes, K., Tjotta, S., 1994. Productivity differences in multiple output industries. Journal of Productivity Analysis 5, 23–43.
Salvanes, K., Tjotta, S., 1998. A test for natural monopoly with application to Norwegian electricity distribution. Review of Industrial Organization 13, 669–685.
Sappington, D., 1980. Strategic firm behaviour under a dynamic regulatory adjustment process. Bell Journal of Economics and Management Science 11, 360–372.
Thanassoulis, E., 2001. Introduction to the Theory and Application of Data Envelopment Analysis: A Foundation Text With Integrated Software. Kluwer Academic Publishing, Massachusetts.
Vickers, J., Yarrow, G., 1993. Privatization: An Economic Analysis. The MIT Press, Cambridge, MA.
Yatchew, A., 2000. Scale economies in electricity distribution: a semiparametric analysis. Journal of Applied Econometrics 15, 187–210.