

Competition and Regulation in Network Industries

Offprint



intersentia

CRNI

TOTAL COST EFFICIENCY ANALYSIS FOR REGULATORY PURPOSES: STATEMENT OF THE PROBLEM AND TWO EUROPEAN CASE STUDIES

VIRENDRA AJODHIA, TARJEI KRISTIANSEN, KONSTANTIN PETROV*
and GIAN CARLO SCARSI**

Abstract

This paper describes the issue economic regulators face when assessing the relative performance of network utilities by means of economic benchmarking. Using examples from two European regulatory agencies, we discuss how total expenditure (totex) benchmarking can achieve more consistent outcomes than building-blocks benchmarking, while generating a new set of practical issues which will make this regulatory tool feasible only under a specific set of (mainly country-specific) circumstances.

Keywords: total cost efficiency analysis, benchmarking, regulation, electricity networks, Data Envelopment Analysis.

JEL Code: L51.

* KEMA GmbH, Kurt Schumacher Str. 8, D-53113 Bonn, Germany Virendra.Ajodhia@Kema.com, Tarjei.Kristiansen@Kema.com, Konstantin.Petrov@Kema.com. Tel. +49 228 446 900, Fax +49 228 446 9099.

** Corresponding author (e-mail preferred correspondence). All co-authors are with KEMA. Virendra Ajodhia is also with the University of Delft, The Netherlands. This paper only reflects the authors' personal views and not those of KEMA. We would like to warmly thank Michael Boxall for valuable discussion inputs on many of the topics treated in this paper. Nonetheless, any mistakes and inaccuracies remain our own. KEMA Ltd, Regent's Place, 338 Euston Road, London NW1 3BT, United Kingdom. GianCarlo.Scarsi@Kema.com Tel. +44 870 351 8900, Fax +44 870 351 8903.

1. STATEMENT OF THE PROBLEM

1.1. PRICE CAP REGULATION

Regulatory authorities dealing with utility sectors are faced with the issue of regulating price and quality levels for natural monopolies. Not all of the production subsets that are typical of a utility sector are naturally monopolistic. Indeed, technology and demand conditions have made large fractions of the utility realm relatively competitive today, although different conditions exist in traditional industries such as electricity, gas, and water as opposed to – say – telecoms and media.

Regulating electricity and water networks, for instance, still requires a high measure of regulatory judgement as to which techniques and methods one should use to make up for the substantial lack of direct technical efficiency and cost comparability between decision units. As a result of the existence of natural monopolies in the fields of transmission and distribution of electricity and gas, regulators have resorted to various forms of either partial or total relative efficiency analysis – and even to pure yardstick competition in some cases. We will refer to this practice by the usual – albeit not entirely correct – name of “economic benchmarking”.

1.2. TOTEX BENCHMARKING

Economic benchmarking can be either carried out by looking at technical variables only (for instance, the length of distribution networks as a proxy for capital and the number of employees as a proxy for labour) or by examining different cost levels. Cost-based benchmarking is more popular with economic regulators because it better fits the regulatory concepts of “efficient allowed cost” and “regulatory revenue allowance” that are typical of incentive-based regulation as applied by a number of European, Australasian, and South American countries.

Cost-based benchmarking can take place on the basis of either partial or total cost. Under the total cost or ‘totex’ approach, the regulator does not differentiate between opex and capex: on the contrary, it sets the X factor on the basis on the sum of these, i.e. on the basis of total cost (totex). In practical terms, this means that the regulator usually does not consider investment projections explicitly.¹ Instead, it performs a benchmarking analysis of actually incurred levels of totex. The resulting efficiency scores, then, are an input to future allowed totex levels.

¹ In this case the regulator looks at existing total cost, judges efficiency on the basis of such cost and applies the results for the future. Future investments are not checked ex-ante, but rather ex-post via repeated benchmarking. The practical application of this approach, however, is usually combined with the usage of flexible revenue cap regimes that may include investment allowances – either explicitly or through observed changes in selected revenue drivers.

The efficiency incentives typical of the totex approach stem from the fact that during each regulatory period, the X factor is set on the basis of the performance achieved in previous years. If the firm manages to increase productivity, its efficiency score will be higher in future periods and, consequently, its X factor will be lower. This is an important divergence with the building blocks approach, whereby assessment issues about capex projections might hinder the correct determination of efficient expenditure levels.

The ideal solution for regulators would be to avoid any trade-offs between separate components of cost, i.e. opex and capex, and to analyse relative efficiency levels of network-based utilities by just looking at total cost levels, after allowing for structural regional differences and other identifiable peculiarities.

However, in the real world did only a handful of regulators manage to perform total expenditure (totex) benchmarking as opposed to other, more partial forms of this exercise. The British energy regulator Ofgem has traditionally preferred the so-called 'building-blocks'² approach whereby cost levels undergo separate scrutiny according to their supposed nature of either capital (capex) or operating (opex) expenditure. Other regulators, however, have taken a different view. Having recognised the 'double jeopardy' problem stemming from building-blocks benchmarking – which results from the allocative and accounting trade-offs between capex and opex – such regulators have moved towards the use of totex benchmarking for the sake of setting revenue requirements, with a view – at least in one case – to applying straight yardstick competition to network operators. Finally, the application of totex-based benchmarking might be preferred for practical reasons whereby a large number of network service providers (for instance in Norway, Austria and Germany) exist, and the regulator is not in the position to check the investment plans of each individual company *ex ante*.

This problem is particularly relevant given that investment benchmarking is a difficult exercise. Investment is, among other things, driven by developments in demand and asset replacement considerations. These factors can be different from firm to firm, and tend to distort comparative analysis. Furthermore, investment is typically lumpy although firms usually have some degree of flexibility in timing their choices. A firm may either decide to postpone investment or to bring it forward. For this reason, directly assessing the desired investment level by means of comparative analysis is problematic – even more so, when the quality dimension is taken into account. Nevertheless, capital cost generally forms a substantial fraction of the firm's

² In the building blocks model, the net present value of capital cost at any point in time is just equal to the level of the regulatory asset base at that time. As a result, the power of incentives for cost reduction on capital expenditure will depend on the responsiveness of the regulatory asset base to the capital expenditure out-turn of the regulated firm. If the regulatory asset base is adjusted *ex post* by an amount equal to the actual capital costs incurred by the regulated firm, then the regulated firm will have a very weak incentive – if any at all – to economise on capital expenditure.

total cost, and ensuring that investment is undertaken efficiently will normally generate significant benefits to society as a whole.

1.3. PAPER OUTLINE

What follows is an overview of two totex benchmarking approaches that try and overcome the possible cost misallocation and allocative inefficiency bias implied by 'waterproof' opex and capex benchmarking, and the problems entailed by the subsequent construction of an infeasible total cost frontier obtained as the sum of the best feasible post-benchmarking opex and the best feasible post-benchmarking capex ('double jeopardy').

The first approach, adopted in Norway by the Norwegian Water and Energy Directorate (NVE), starts from conventional cost analysis, and goes forward to develop a total cost benchmarking framework with control factors that arrives at the determination of efficiency and revenue requirements for electricity distribution utilities. The second approach, developed in the Netherlands by the Dutch energy regulator DTe, is even more ambitious in that it turns total cost benchmarking into an almost completely mechanistic transposition tool to arrive at a semi-automatic yardstick competition regime. Both approaches were originally intended for electricity distribution utilities, and use Data Envelopment Analysis (DEA) as the main benchmarking technique.

Firstly, the benchmarking technique used by both countries is quickly described. Secondly, the two totex benchmarking approaches are discussed in turn, and an evaluation of the totex approach is made. A final concluding section draws up a few considerations for future research.

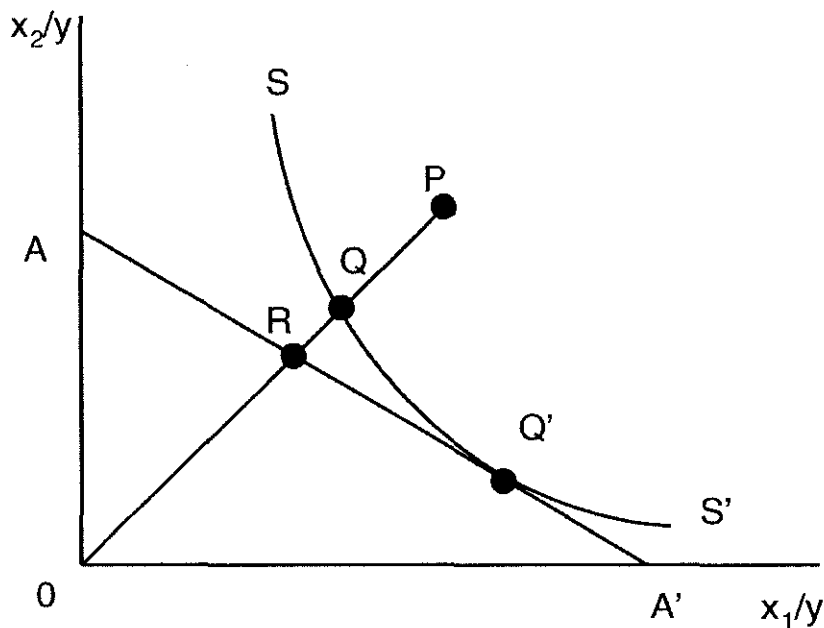
2. DATA ENVELOPMENT ANALYSIS (DEA)

DEA is a non-parametric mathematical programming approach to productivity frontier estimation. The general idea behind DEA is the measurement of a firm's productivity performance by observing its distance to the productivity frontier, which is constructed on the basis of the best performing firms ('peers') in the given data sample. According to Coelli et al. (1998), the foundation for DEA was laid by Farrell (1957), who in turn drew upon the work of Debreu (1951) and Koopmans (1951).³ Farrell (1957) defined a simple measure of productive efficiency that could account for multiple inputs, and could easily be generalised to multiple outputs. He claimed that the productivity of a firm consists of two components:

³ An overview of DEA is in Seiford and Thrall (1990). Two basic DEA models to which most applied papers still refer are those by Charnes, Cooper, and Rhodes (1978) for constant returns to scale (CRS) DEA, and by Banker, Charnes, and Cooper (1984) for variable returns to scale (VRS) DEA.

- firstly, technical efficiency, which reflects the ability of a firm to obtain maximum output from a given set of inputs, and
- secondly, allocative efficiency, which reflects the ability of a firm to use different inputs in optimal proportions, given their respective prices. These two measures are then combined to provide a measure of total productive efficiency.

Figure 1. Technical and allocative inefficiency



Source: Coelli et al. (1998)

Farrell (1957) illustrated his ideas by using a simple example involving firms that utilise two inputs (x_1 and x_2) to produce a single output (y), under the assumption of constant returns to scale, i.e. output changes are proportional to input changes for all levels of input. The unit isoquant of the fully efficient firm is represented by SS' in Figure 1. This efficient isoquant represents the minimum amounts of inputs x_1 and x_2 needed to produce one unit of output y . Under the assumption that the efficient isoquant is known, the technical efficiency of a given firm can be calculated. A given inefficient firm (i.e. one that is not located on the curve SS') uses a certain quantity of inputs, defined by point P , to produce a unit of output. Technical inefficiency for that firm is represented by the distance QP , i.e. by the amount by which all inputs could be proportionally reduced without a reduction in output. This is usually expressed by the ratio QP/OP , which represents the percentage by which all inputs could be reduced. The technical efficiency level (TE) of a firm is most commonly measured by the ratio:

$$TE = OQ/OP = 1 - QP/OP \quad (1)$$

This ratio takes a value between zero and one, and provides an indication of the degree of technical (in)efficiency of the firm. A value of one indicates that the firm is fully technically efficient. For example, point Q is technically efficient because it lies on the efficient unit isoquant.

Notice further that all DEA measures are bound by zero and one, and that overall (Farrell) efficiency has the property of multiplicative separability into input-allocative and technical efficiency.

The basic idea behind DEA, as proposed by Charnes et al. (1978), is to solve the efficiency score program for each firm based on a constrained linear formulation. We refer the interested reader to this paper for a mathematical formulation of the DEA linear program, and for further decomposition of efficiency measures into allocative and other components.

The Norwegian Totex Benchmarking Strategy

Institutional Background

The Norwegian Water Resources and Energy Directorate (NVE) is the power industry regulator in Norway. The NVE is a directorate under the Ministry of Petroleum and Energy, with responsibility for managing the country's water and non-fossil energy resources. NVE's mandate is to ensure integrated and environmentally friendly management of the country's watercourses, to promote efficient energy markets and cost-effective energy systems, and to work towards the achievement of a more efficient use of energy. NVE is also responsible for reducing damage caused by floods and erosion along rivers and streams.

Network regulation in Norway started out as a rate-of-return scheme during the first price control period (1992-1996). However, the deficiencies resulting from this approach were soon recognised. The main issues turned out to be the inefficiency caused by guaranteed ex-post cost recovery and the weak incentives for productivity improvement. These considerations became the major reason leading to the replacement of the existing price control framework by a new incentive-based scheme in 1997. The current Norwegian regulatory system is an ex ante regulation method based on incentive regulation and revenue caps. Through efficiency incentives, NVE strives to encourage network owners to reduce costs and improve their efficiency. Under the new system, network owners are no longer guaranteed full cost recovery. By establishing a system whereby each network owner is allowed to receive pre-determined maximum revenue, the allowed profit will in principle coincide with the difference between allowed revenue and actual cost. Allowed revenue requirements should cover the networks' total cost: operation and maintenance; capital cost in the form of depreciation and return on capital invested; network losses; and profit taxes.

The regulatory model of Norway is based on revenue cap regulation, supplemented by benchmarking and profit-sharing mechanisms. Initial revenue caps were determined on the basis of regulatory accounts from 1994 and 1995. In 2002, NVE reset the price control for networks leaving the general logic of the revenue cap from the first regulatory period (1997–2001) untouched, however adjusting some of its components and pursuing the improvement of the regulatory cap's properties.

Price Regulation

Norway's revenue cap regulation for electricity transmission and distribution contains elements of different regulatory mechanisms. It consists of cost-plus regulation with a time lag, and benchmarking plays a crucial role in determining efficiency requirements. The revenue cap is determined by:

- The revenue cap from the preceding year, or primarily the cost in the first year of the regulation period, plus a standard return on capital for the same year;
- An expected efficiency improvement parameter, benchmarking-based; and
- An annual correction factor intended to provide an additional revenue as a function of pre-specified cost drivers.

The fact that the revenue cap is affected by these factors makes it possible for the grid owner to influence its return – not only by decreasing costs, but also by operating and maintaining the grid in such a way to ease the revenue cap.

The initial revenue values consist of the average operating and maintenance costs, depreciation, return⁴ on invested capital (book value + 1% to allow for working capital), and average grid losses. The spot market price for power is used to assess the value of grid losses.

As said, in the first revenue cap period (1997–2001), the initial revenue caps were determined on the basis of grid company accounts from 1994 and 1995, according to the following formula:

$$RC_e = OPEX + DEP + r * RAB + L$$

where RC_e is the initial revenue cap determined by operating and maintenance costs (OPEX), depreciation (DEP), return on invested capital (r) on the asset base (RAB), and costs associated with energy losses (L). We also have:

$L = L_{MWh}$, where:

L_{MWh} denotes losses in MWh, and P is the average system price of energy during the whole year as gleaned from the NordPool spot market.

⁴ NVE annually sets a reference interest rate based on the long-term risk-free interest rate plus a risk premium. The reference interest rate, rn , is the basis for the return. The basis for the risk free interest rate is the ST4X index from the Oslo Stock Exchange. The risk premium is valued at 2%.

The dynamic time adjustment of the allowed revenue for the grids was based on the following formula:

$$RC_n = ((RC_{n-1} - RC_{n-1} \cdot \left(\frac{CPI_n}{CPI_{n-1}}\right) + L_{MWh} \cdot P_n) \cdot (1 - X) \cdot 1 + \frac{\Delta E_{a,n}}{2},$$

where:

- RC is the starting revenue cap (revenue requirement based on cost);
- CPI is the consumer price index;
- X is the efficiency requirement calculated at the beginning of the regulatory period by means of Data Envelopment Analysis (DEA);
- ΔE is the percentage increase in transported energy on a year-on-year basis; and
- n is the yearly time index.

The initial revenue is annually adjusted for inflation, the required efficiency increase, and by the term $(1 + 0.5 \cdot \Delta E_{a,n})$. The latter is designed to provide an additional revenue to the regulated grid companies intended to contribute towards the additional OPEX and CAPEX being incurred as a result of the increasing volume of transported energy. The anticipated efficiency improvement includes:

- An individual efficiency increase component – measured via DEA on inter-company comparisons; and
- A general efficiency increase component – imposed exogenously by NVE and reflecting the general technological improvement in the industry.

The current revenue cap scheme (2002-2006), with initial cost values taken from 1996-1999, is slightly different from the above description. Amongst other things, the last term in the formula was removed and replaced by a new term (*Adjust*) that adjusts for new investment, whereas the increase in transported energy was supplemented by a second driver, namely the relative increase in the number of buildings in each distribution area:

$$RC_n = ((OPEX + DEP + RAB \cdot r_n) \cdot \left(\frac{CPI_n}{CPI_{2000}}\right) + L_{MWh} \cdot P_n) \cdot (1 - X) + Adjust_n^i,$$

whereby the notation is as above. The Norwegian approach relies on the incentive to raise efficiency in the regulatory period while keeping any interim efficiency gains. The regulatory revenue re-setting is based on actual cost, including checks for any deviation between the prescribed revenue path and actual performance. In addition, the regulated grid providers are exposed to repeated benchmarking that is aimed to eliminate inefficiencies. In the current regulation period (ending 2006), the grid owner will not be able to influence the revenue cap directly. The revenue path is decoupled

from actual cost via the application of the regulatory formula.⁵

If we assumed that last year's cost level is the only basis for determining the revenue cap in the current period, we would impose no cost-reducing incentives on the regulatees. What NVE does instead is to benchmark the opening level of costs via DEA, and to impose an efficient totex level objective upon all regulated companies throughout the duration of the regulatory period.

The Efficiency Requirement

The efficiency requirement contained in the formula (the X term) is based on a DEA comparison of all distribution companies (there are more than 150 in Norway, albeit now decreasing). As regards asset values, both book asset values and replacement asset values are considered as part of the total-cost DEA runs, which are then computed twice. Companies are given the 'benefit of the doubt' in that the most favourable DEA scores will be used for their revenue requirement calculations after comparing the efficiency score series from the two DEA runs. The relationship between the efficiency requirement and efficiency measurements based on DEA is softened in such a way that the individual cost improvement requirement will never exceed 3% annually for any distribution company reporting a DEA cost efficiency score of 70% or lower ('efficiency flooring'). Formally, the efficiency requirement target X is given by the formula:

$$X_D = 1 - (1 - (1 - EFF) \cdot 0.3824)^{1/4},$$

where EFF is the totex efficiency level for any given company as calculated via DEA, which is – as said – floored at 0.70 for all distribution companies with a reported 'raw' DEA cost efficiency score of less than such amount. For a grid owner with regulatory cost efficiency in the floored 70-100% interval then, the formula will mean that 38.24% of the individual inefficiency in the distribution grid must be recovered over the regulatory time span of 4 years. Any residual inefficiency will be carried forward to the following regulatory period.

Dealing with New Investment

There is an adjustment term in the revenue cap formula that exceeds the compensation implicit in the incentive to 'beat' the regulatory benchmarking-based cost target. On the assumption that new investment may be caused by an objective need resulting from changes in certain cost drivers, as well as being driven by safety and system security reasons, the investment revenue adjustment element in the

⁵ Differently from the UK, the revenue formula does not use any building blocks and does not include any ex-ante expenditure projections.

revenue cap formula is intended to give the companies suitable compensation for expanding the grid. However, it is important that the adjustment term does not favour unnecessary and/or gold-plated new investment. In other words, the adjustment term should not incentivise the companies to influence their own revenues through uneconomic actions.

New investments involve capital costs such as depreciation and return on invested capital. The majority of such costs is already taken into account by updating the cost base for the revenue cap periodically. Cost recovery is, however, delayed in time because updates do not occur continuously. This entails that the net present value of the implied revenues is lower than what would be necessary to cover new capital cost incurred today. The purpose of the investment adjustment term in the revenue cap is to provide continuity in terms of investment recovery. In addition, new investment may have an impact on operation and maintenance cost, grid losses, and undelivered energy. Such (arguably positive) changes will not result in changed revenue caps for the firm until the next regulatory review, and must therefore be assessed when determining the level of the ongoing adjustment term. The share of capital costs associated with new investment that the adjustment term is supposed to cover depends on:

- The real timing of investments in relation to the four-year update timetable for totex;
- The life time of the investments;
- Future inflation;
- Future efficiency requirements; and
- The return on capital (discount factor) for the grid owners.

NVE has calculated that between 64.8% and 94.5% of capital costs are already covered through the four-year totex revenue cap. The adjustment term shall therefore compensate for between 5.5% (100% minus 94.5%) and 35.2% (100% minus 64.8%) of the capital cost associated with the new investment.

3. THE DUTCH YARDSTICK COMPETITION EXPERIMENT

3.1. INSTITUTIONAL BACKGROUND

Before liberalisation (occurred in the late Nineties), electricity tariffs in the Netherlands were set by a system that closely resembled cost-plus. Under this system, tariffs were primarily based on observed costs, plus a reasonable rate of return. Although much less legalistic and explicit than in the US, the generally observed weaknesses of the traditional rate-of-return regulatory approach also applied to the Netherlands. The 1998 Electricity Act introduced a completely new approach towards price regulation. Currently, tariffs for distribution network use are regulated on the

basis of a price-cap system. Tariff levels are annually adjusted by $CPI - X$, whereby CPI is the consumer price index and the so-called X factor is the regulator's estimate of future efficiency improvements.

For the first regulatory period, from 2001 to 2003, regulator DTe's price capping scheme could be characterised as an *ex ante-cum*-benchmarking approach. Tariffs were set beforehand for a number of years on the basis of regulatory judgement about future productivity improvements as disclosed through benchmarking. For both electricity and gas network monopolies, price-cap systems have been used as a basis for setting tariffs. The strategy for setting the X factors was to drive firms towards jointly achievable efficiency levels in order to create a level playing field, so that yardstick competition could be subsequently introduced in the second regulatory period starting 2004 (DTe 2002).

3.2. PRICE REGULATION

In the second incentive-based regulatory period (2004-2006), DTe decided to move towards fully-fledged yardstick regulation of price and quality. However, the application of yardstick competition by DTe is different from the classic textbook approach (Shleifer, 1985). According to Shleifer's original proposal, price is set on the basis of average ($n - 1$) cost. In the Dutch case, prices are simply *adjusted* by an X factor reflecting the change in the total factor productivity of frontier-located firms. Thus, in the yardstick formula prices are based on relative changes in productivity rather than on absolute costs.

For each firm, an annual allowed revenue is determined. Based on this, tariffs are set such that allowed revenue equals the sum of revenues generated by these tariffs and the actual volumes as observed back in 2000. These volumes, y_{2000} , act as so-called "norm volumes". The calculation is made as a tariff basket norm whereby, for each year, prices are set such that company i will be allowed to earn the following revenue:

$$RC_i = \sum P_{j,t} y_{2000, j, t}$$

RC is the revenue cap and results from the yardstick scheme; p_j stands for the j^{th} tariff/price component. This formula is the starting point for setting annual tariffs.

RC comes from the opening, total-cost benchmarking and is set on the basis of the yardstick scheme, following the formula below:

$$RC = (1 - X)^t RC_{2003},$$

where t runs through the regulatory period's years 2004-2006.

3.3. THE EFFICIENCY REQUIREMENT

So far, the Dutch system could comfortably be described as a straightforward tariff-basket system. The main issue that sets this system apart from others, however, is the way in which the X (efficiency requirement) factor is set. X is set according to the following formula:

$$(1 - X_j)^3 = EFF_i \cdot (1 - FS)^3.$$

Here, EFF_i is the efficiency score for each particular company. This efficiency score is based on the 2000 totex Data Envelopment Analysis study (the latest one available in the Netherlands), putting together capital and operating cost, and is adjusted for possible efficiency improvements made in the 2000-2003 period.

The factor FS is the “frontier shift”, measured on the basis of the change in total factor productivity for frontier companies (those with an initial or corrected efficiency score of $EFF = 100\%$).

Note that the X factor is firm-specific, but the FS one is the same for all companies. Thus, each company has its own X factor, which consists of two separate components:

- A general target (FS), which is the same for all companies and is based on the change in total factor productivity as observed at the frontier throughout 2004-2006; and
- An individual target (EFF_i) which is set individually by company, and aims at removing initial efficiency differences across companies.

The general idea behind the Dutch yardstick system is that prices are adjusted on the basis of realised changes in general productivity (i.e. the frontier shift). This is done by measuring the change in the frontier’s position (as reflected by the FS factor). FS is only determined by those companies that are 100% efficient on a total (not just operating) cost basis, i.e. those that either initially have a total cost DEA score of 100% or have otherwise caught up with the frontier in the meantime.

The reason for splitting the X factor into two components stems from the need to recognise that those companies that are initially inefficient should not be included in the calculation of the frontier shift. Thus, companies that have less than 100% efficiency will be given two targets, i.e. (1) catching up with the frontier; and (2) shifting together with the frontier.

After the second regulatory period, the value of EFF will be set to one by default. Therefore, DTe will assume that all companies are equally efficient (level playing field) and that no further adjustment is needed to the yardstick scheme. At that stage, the X factor will be set on the basis of FS only.

3.4. YARDSTICK-BASED BENCHMARKING

The value of EFF_i is known, and directly follows, from the 2000 totex DEA study bundling opex and capex efficiency analysis together. The value of FS is not known at the start of the period, but can only be measured once the productivity improvements have been achieved ex post. Therefore, DTe will have to forecast FS at the start of the regulatory period. At the end of the period, DTe will measure the “true” FS and will correct for any differences – including interest – in the subsequent regulatory period.⁶

On top of the above formulas, a CPI adjustment is applied (CPI-X). Note that, for the purpose of measuring productivity changes and making any appropriate ex post corrections, costs need to be restated to a common base year.

It is worth noting that, at the start of the third regulatory period, DTe will not yet have full information about cost as incurred in the last year of the second period. Therefore, the comparison between forecast and actual improvements will exclude this last year. That is, FS will be based on changes measured in the period 2003–2005. As already mentioned, FS is derived from the relative change in productivity at the frontier. Here, productivity is more precisely intended as Total Factor Productivity, i.e. it is defined in terms of an index number consisting of a ratio of weighted outputs to weighted inputs.

As regards inputs, DTe uses the following items to make up total cost (TOTEX):

- Operating cost (OPEX);
- Standardised depreciation (DEP); and
- Return on standardised book values (WACC times the Regulatory Asset Base or RAB).

Inputs are weighted equally, i.e. the single input in the comparative efficiency exercise is simply given by:

$$TOTEX = OPEX + DEP + [WACC * RAB].$$

As an output, DTe uses a so-called “proxy” which is defined as the weighted sum of sold quantities (measured by the number of customers, kWh, kW, etc.) for all customer classes.

The weighting is made on the basis of predefined factors which are set by DTe. The weighting factors correspond to an average tariff basket for the industry.

In the totex efficiency analysis, a company will become efficient if its costs per proxy output in 2002 are less than, or equal to, the projected/target efficient level (which is the efficient, or frontier, cost level obtained under totex DEA). The general view of the regulator is that a company will be labelled a ‘frontier’ company if it manages to achieve the efficiency improvement required by its DEA score in 2000 by

⁶ DTe’s forecast for FS during the first regulatory period was 1.5% per annum.

the end of the regulatory period. Any super-efficiency achievements will be pocketed, in the form of extra profit, by the over-achievers.

Once a company reaches the frontier, it will be considered fully efficient in the future 'by default'. That is, all firms with a DEA score of 100% in the 2000 benchmark will 'by definition' stay on the frontier – irrespective of their achieved performance afterwards. This is a courageous, and potentially faulty, policy assumption which might turn out to be wrong *ex post* because of efficiency leap-frogging.

4. EVALUATION

Under the totex approach, the problem of investment assessment may be effectively bypassed. Furthermore, as the totex approach does not distinguish between opex and capex, the firm (as well as society) may also achieve efficiency gains by trading-off between labour and capital inputs. Under the totex approach, the regulator does not need to develop a view as to whether a given investment proposal should be allowed or not. Rather, the regulator will consider the actual total cost (including investment) incurred by the firm, and will set the X factor based on the analysis of such costs.

4.1 CAPEX MEASUREMENT

The measurement of capex is traditionally problematic. Investments are typically undertaken at different time intervals and tend to considerably vary in size. Investment lumpiness might be characterised by substantial fluctuations in cash spending from year to year, which could lead to misleading results in the benchmarking. Averaging capex spending for a number of years can partially smooth out the figures, but will not completely account for differences, in particular when companies turn out to be at different stages of their investment cycles.

A possibly superior alternative might be to represent capex as an estimate of annual capital consumption. According to this approach, a stream of annual investment figures is converted into a stock of assets (on which a return is earned), and into a stream of annual depreciation figures. In this case, annual capital consumption is measured as the sum of the capital cost's components: return on assets and depreciation. This method replicates the canonical regulatory accounting scheme used to establish allowed revenues, and paves the way for total cost benchmarking of a top-down type using either econometric or linear programming techniques. However, network age and investment cycles are not properly reflected in this model and may well distort the credibility of results.

NVE has used both book and replacement values as a basis to establish the annual

capital consumption⁷ in its benchmarking estimates, and has tended to provide the regulated firms with the 'benefit of the doubt' by recognising whichever value was more advantageous to them in terms of cost efficiency assessments (Bjørndal, M. and K. Jørnsten, 2003).

4.2. THE LONG-TERM NATURE OF INVESTMENT

There is an important issue to raise about the totex approach, which is related to the long-term nature of investment. Capex (depreciation and return) is spread over a number of years and therefore, ideally, the benchmarking exercise would need to track it down over a long enough period of time rather than taking a simple snapshot of (however standardised) costs from a single financial year.

Costs in a given year can be strongly influenced by the timing of investment. In the long run, similar firms should face the same level of depreciation costs, but even when other things are kept equal, the firms will probably locate at different investment stages in the short run. As a result, *ceteris paribus*, one firm might have high depreciation costs in the early years, and relatively low depreciation costs in later years.⁸ If benchmarking considered only a single year without standardisation, two otherwise identical firms would turn out to have potentially high efficiency divergences depending on the timing of benchmarking vis-à-vis the timing of investment.

As the example from the Table below shows, costs in a given year can be strongly influenced by investment timing. In this simple example, firms A1 and B1 both invest an amount of 400. In the long run, these firms will face the same level of depreciation cost. However, their investment timing differs. Firm A1 invests primarily in the last year while for firm B1, most investments are conducted in the first year. As can be observed, the effect of this is that firm B1 will have high depreciation costs in the early years, and relatively low depreciation costs in later years. If the benchmarking analysis considered only one single year, say the second one, firm A1 would turn out to be very efficient as it would have much lower costs (50) as compared with firm B1 (88). The reverse would apply if the benchmarking analysis were conducted later in time, e.g. in the last year.

This example demonstrates the importance of considering multiple years in a totex benchmarking analysis. Multi-year analysis, however, would make the benchmarking exercise much more data demanding, and therefore less practical. Consideration should also be given to the fact that the analysis may be hampered because of differences in cost measurement resulting from different accounting conventions in

⁷ In the case of book values, annual capital consumption is defined as the sum of depreciation and return on assets. In the case of replacement values, annual capital consumption is defined as an annuity being calculated by using the allowed rate of return and the average network asset life.

⁸ In the interest of simplicity, only depreciation costs are compared. Similar effects would, however, apply to returns.

the treatment of capital cost. Consider the above example once more. Firms A1 and A2 (or B1 and B2) both have the same investment pattern, but use different depreciation periods. Because of this, their depreciation costs measured in the same year tend to be different. Firm A1, which uses a depreciation period of four years, has lower costs in the earlier years than firm A2 – the latter using a shorter period of two years. Conversely, in the later years, firm A1 still incurs depreciation costs while firm A2 would have already depreciated all of its assets.

Although the examples provided here are simplified, they illustrate the basic problem of how ignoring the long-term nature of investment can distort benchmarking results. Even though in the long run firms will have invested at similar levels, their costs would still fluctuate from year to year, reflecting differences in investment timing and depreciation policies. Including multiple years in the analysis could help, but will not solve this issue. It would also make the analysis more data demanding. This is particularly true when the firms considered in the analysis have been using different accounting conventions throughout. Performing a backwards calculation of book and depreciation values (standardisation) could eliminate the monetary effects resulting from differences in depreciation policies, but not those in investment timing. However, there may be a problem in obtaining such historical data, in particular given the relative long lifetimes of assets in the electricity distribution business.⁹

Table 1. Simplified example of the impact of different depreciation policies and investment timing

All firms invest the same amount over a period of three years and use straight-line depreciation, but differ in the timing of these investments and in the depreciation period. Although in the long run depreciation costs are the same, annual depreciation will vary considerably.

Firm A1 (depreciates in 4 years)					
Year	Investment (mIn. EUR)	Depreciation (mIn. EUR) for investments in year:			Depreciation Costs (mIn. EUR)
		1	2	3	
1	100	25			25
2	100	25	25		50
3	200	25	25	50	100
4		25	25	50	100
5			25	50	75
6				50	50
Total	400	100	100	200	400

Firm B1 (depreciates in 4 years)					
Year	Investment (mIn. EUR)	Depreciation (mIn. EUR) for investments in year:			Depreciation Costs (mIn. EUR)
		1	2	3	
1	300	75			75
2	50	75	13		88
3	50	75	13	13	100
4		75	13	13	100
5			13	13	25
6				13	13
Total	400	300	50	50	400

⁹ In the Netherlands for example, lack of suitable data made any accurate correction of capital costs difficult. This was an important factor leading regulated firms to reject the benchmarking analysis conducted by the Dutch regulator (Ajodhia et al., 2003).

Total Cost Efficiency Analysis for Regulatory Purposes

Firm A2 (depreciates in 2 years)					
Year	Investment (mln. EUR)	Depreciation (mln. EUR) for investments in year:			Depreciation Costs (mln. EUR)
		1	2	3	
1	100	50			50
2	100	50	50		100
3	200		50	100	150
4				100	100
5					-
6					-
Total	400	100	100	200	400

Firm B2 (depreciates in 2 years)					
Year	Investment (mln. EUR)	Depreciation (mln. EUR) for investments in year:			Depreciation Costs (mln. EUR)
		1	2	3	
1	300	150			150
2	50	150	25		175
3	50		25	25	50
4				25	25
5					-
6					-
Total	400	300	50	50	400

Finally, an annuity approach that converts the investment into a constant annual capital consumption can be used to eliminate the effect of investment cycles. Ideally, the annuity should be calculated for each investment item within each asset group in the past, and then added back together in the year when benchmarking takes place. To convert investments into annuities, the regulator can use the allowed rate of return and asset life (by asset group) as applied to establish the revenue requirements.

Ofgem in the UK has not applied yet any total cost analysis starting from standardised capital expenditure data. An alternative approach for the UK regulator would be to embark on multi-year benchmarking analysis as long as existing data sets are made homogeneous. Time-series analysis might at least mitigate the problem of capital cost estimation, and could open new possibilities for total cost benchmarking – even without having to adopt the radical views of the Dutch energy watchdog.

4.3. CONTROLLABILITY

Whatever is directly under managerial control in the short run is deemed as controllable. Controllable costs include for example the costs of maintenance, personnel, office rents, transportation, etc. In the long run, ideally, investments are also controllable as the firm is free to determine its investment policy and the proper mix between capex and opex. This is, in particular, important for the application of the totex approach whereby controllability is extended to capital cost components.

In the Netherlands, a selected number of cost items are considered to be controllable. These include the cost of network losses, network congestion, and reactive power. These costs are in principle allowed on the basis of actually incurred cost levels. Ex ante, a projection of cost levels is made which is corrected for ex post out-turns, on the basis of actually incurred costs. Thus, there is an ongoing correction in tariff levels, which takes place on an annual basis.

In Norway, there are fixed operating costs that are independent of load. Such operating costs are sometimes deemed as controllable items, and sometimes not (e.g., those relating to environmental conditions not modifiable by the utility). Capital expenses, once recognised, are by definition controllable. Investment items being carried over from the past, given regulatory and ownership continuity that are typical of Norway,

are normally considered as falling within full responsibility of the utility, unless a merger had taken place in the meantime, in which case NVE will have to redeploy its regulatory accounting weapons and embark on supplementary, off-model analysis to define on which (case-by-case if needed) basis any overlapping capital cost items from mergers should be allowed as part of the post-merger regulatory benchmarking framework. At merger point, cost duplication might become an issue (whether or not fully offset by economies of scale and/or scope) and NVE is entitled, as a result, to reclassify costs as either controllable or non-controllable, and to either allow or disallow such cost items. Finally, in Norway variable operating costs that are directly related to the distribution activity as reflected by connections management, grid maintenance, serving load, and other similar distribution cost drivers, enter the benchmarking exercise directly, and as such are considered by NVE as falling fully within managerial control.

Moreover, as explained earlier in this paper, benchmarking analysis is usually backward-looking and therefore it implies (when applied to totex) controllability over investments undertaken in the past. However, not every investment and, correspondingly, no annual capital measurement (return plus depreciation or annuity) is in reality truly controllable. There might be some cost items that cannot be considered as being within the company's scope of control both in the short *and* long run. Therefore, such non-controllable costs should not be included in the benchmarking analysis but, on the other hand, they should be passed through. While the differences in the companies' operational environment and the corresponding objective cost differences can be reasonably considered in the model specification via additional output variables, there are still two additional factors that should be accounted for: externally imposed requirements, and historically inherited constraints.

Typical non-controllable cost items stemming from externally imposed requirements are:

- force-majeure related costs;
- state security-related costs;
- environmental compliance costs;
- legislation and regulation compliance costs; and
- costs stemming from other externally-imposed requirements (e.g. rents and leases, rights of way etc.).

All these costs may be relevant for the provision of the regulated network service, but should be excluded from benchmarking.

The second, more difficult factor deals with the question of whether – and to what extent – costs incurred in the past could be considered controllable today – due to the reason that decisions made in the past by a different management team cannot always be reversed by the company's current management. This argument becomes stronger if equity control has changed in the meantime, and the company's shareholders today are different from those controlling the company in the past. Finally, based on

currently available knowledge, it may be difficult to distinguish clearly whether past decisions were correct or not, as such decisions only reflected the knowledge, information, and technology available at that time.

To account for such historical pre-conditions, the regulator may decide to exclude the assets acquired before a certain threshold date (e.g., to only consider the assets acquired in the last 20 years or so). While this approach appears reasonable and practicable, it may raise questions as to whether the choice of the appropriate backward-looking window period is fair to all of the benchmarked companies. For example, companies that had relatively young networks at the start of the window period, and consequently spent less in the following years, will appear to be more efficient than the companies that inherited older networks before the threshold date, and therefore had to spend more afterwards.

4.4. STATISTICAL EFFICIENCY VERSUS ACCOUNTING TRADE-OFFS

It might be argued that estimating cost functions for regulatory purposes based upon total expenditure could result in some loss of estimation accuracy as opposed to estimating cost on a separate basis, i.e. by cost item. However, such loss must be traded off against the opportunity for a regulator to limit accounting gaming (capitalisation policies) by regulated firms, and impose that relative efficiency be assessed on a total cost basis, which would limit both cosmetic and real distortionary cost allocation policies on opex and capex. Based on regulatory cost gaming occurrences as observed in several countries, the benefits of doing away with cost apportionment gaming in regulation would probably outweigh any estimation efficiency compromises on the cost function front as a result of giving up opex vs. capex evaluation in favour of joint 'totex' estimation. Recent surveys by Jamasb et al. (2003, 2004) found out that strategic capitalisation of opex under incentive-based network regulation has indeed been faced by utility regulators. Empirical evidence and instances of this behaviour in the UK are presented by Burns and Davies (1998).

4.5. A COMPARISON BETWEEN TOTEX ANALYSIS AND THE BUILDING BLOCKS APPROACH

Totex analysis for revenue setting purposes in electricity distribution can be characterised as the higher state of incentive regulation as opposed to the building blocks approach, which in its making and more recent versions in a mature regulatory environment, more and more resembles traditional rate-of-return regulation. The table below summarises a structural comparison between the two approaches (the building-blocks version we are considering is the one adopted by Ofgem in the UK; some Australian jurisdictions adopt a similar approach) from a regulatory philosophy standpoint.

We would be comfortable in concluding that, although the Dutch approach is perhaps conceptually superior from a purely economic standpoint, the Norwegian one might be easier to apply in practice for actual regulatory purposes, given its higher flexibility in accommodating investment cost drivers, and its recent turn towards a mixture of benchmarking-based and ideal network-inspired regulation, which definitely constitutes ground for future empirical research.

Methodology	Building Blocks	Totex
Treatment of OPEX	Separately via benchmarking or other efficiency considerations	Integrated in the total cost measure to avoid trade-offs
Treatment of CAPEX	Separately via individual assessment and ex-post adjustment subject to benefit-sharing	Integrated in the total cost measure to avoid trade-offs
Investment allowance	When cliff-edge approaches, the regulator widens ex-ante building block allowance but does not link it to drivers	Investment ex-post benchmarked as part of the totex model (backward-looking), and projected via capital expenditure drivers' analysis
Quality regulation	A separate block (e.g., UK IIP orevolution)	Completely integrated into totex analysis for a fair trade-off against capex
Replacement investment	Allowed for via higher capex allowance, not forward-looking	Allowed inherently in totex via investment drivers' analysis and projected into the future
Sliding scale mechanism	Used ex-post to treat and allocate capital overspend	Not strictly necessary

As shown in the Table grid, totex analysis presents the following strengths as opposed to traditional building blocks:

- It avoids any opex/capex trade-offs which are sometimes devoid of any economic, or even accounting, logic;
- It either benchmarks or statistically analyses investment via the use of investment drivers as opposed to assessing investment plans on a micro basis;
- It enables integrated price/quality regulation;
- It is able to identify investment cost drivers that project replacement investment requirements into the future;
- It can avoid the need for arbitrary regulatory decisions on for the benefit-sharing treatment of capex overspend: under totex, capex is subject to estimation and benchmarking, not to comparisons between different valuations (regulator's vs. companies').

4.6. HOW TO MODIFY THE EXISTING TOTEX APPROACH TO ALLOW FOR CLIFF-EDGE INVESTMENT

The issue of 'cliff-edge' investment requires some extra thinking. Adding specific cost drivers as explanatory variables to the investment regression proposed as part of the Norwegian totex approach makes it possible for firms facing cliff-edge investment needs to properly factor in such needs in the investment requirements once a sufficient allowance is made for both environmental factors and demand growth considerations on the one hand, and for the specific replacement needs of older infrastructure on the other.

It is not possible to evaluate a shift to the totex approach *ex ante* without using real-world data, but it might be worth making totex one of the aspects of a pilot regulatory project focussing on how to overcome the existing limitations of building-blocks inspired network regulation. Regulators might have to take a more consistent and 'compact' view of how to assess Distribution Network Operators' efficiency in the future on an overall basis, taking into account additional drivers of investment expenditure. Such additional drivers – on top of traditional throughput, output, and connection measures – to be considered in relation to extra investment might be the following:

- Regional GDP growth and area building/redevelopment indices;
- Average age of existing assets in the region;
- Average historic total factor productivity of the domestic distribution network industry; and
- Historic network quality indicators from the relevant region and beyond.

Without fully considering regional and national trends in demand, connections, quality, asset vintage, constructions, and building development, a regulator would not normally be able to correctly estimate the actual and prospective efficient level of totex, and to translate it into a revenue requirement figure for electricity distribution companies. Taking a compact and coherent approach to totex and cost-driver evaluation seems to be a superior alternative to the traditional regulatory analysis of individual cost and network quality items.

The need for ensuring that investment is carried out not only to satisfy growing demand and quality requirements, but also to upgrade existing and replace ageing assets, implies that not only should the comparability of totex measurements be ensured, but also that a degree of supplementary analysis might have to be undertaken to accommodate extra investment on top of business-as-usual (like-for-like) requirements.

Supplementary totex analysis implies that the revenue requirements could be augmented by enabling 'on-top' adjustments to allowed totex in consideration of the peculiar regional/national factors resulting from historical lack of replacement and

upgrading investment. Therefore, once total cost analysis is completed and revenue requirements have to be set, the wedge between efficient allowed total cost (following benchmarking analysis) and capital adjustments – as observed, for instance, in Norway – leading to the translation of efficient allowed total cost into revenue allowances, must be bridged by analysing the need for top-up investment as determined by specific company-based and regional capex drivers. The second-stage totex driver analysis approach has the advantage of not doing away with the concept of Regulatory Asset Value, which is relevant to – and valued by – shareholders and financial institutions alike, and – at the same time – it does carry the advantage of accommodating the need for expanding investment in the future as a result of the company's current position in the investment cycle.

For a company which is in a trough of the investment cycle, adding on-top capex driver analysis to the already determined level of efficient total expenditure as a result of totex benchmarking, is an advantage to the extent that it would probably determine a higher level of allowed total expenditure, and build up a more substantial allowance in terms of revenue requirements as the next regulatory review approaches.

5. CONCLUDING REMARKS

Regulation of networks based on total cost benchmarking, as opposed to building-blocks and separate cost assessments, carries an obvious theoretical advantage in terms of creating a level playing field, and ensuring that efficiency frontiers are set by regulators on a fair and achievable basis (hence avoiding the 'double jeopardy' problem). Furthermore, under the totex approach, the problem of investment appraisal as encountered under the more traditional building blocks approach is of lower regulatory priority.

However, the application of regulatory analysis based on total cost benchmarking should be accompanied by a number of considerations to ensure that this kind of analysis is practically feasible, as well as theoretically advantageous.

The totex approach may provide a stable and attractive long-run equilibrium, but may also require some additional considerations, in particular for non-repetitive or one-off investments. The regulator will normally use information on past observed capex cost out-turns in setting the capex forecast for similar projects in future periods. The extent to which past capex cost out-turns provide useful information in setting future capex forecasts obviously depends on how much the capex projects undertaken in the past are similar to those proposed for the future. In the case of capex projects which are repeated regularly (either because the asset involved has a short life or because the firm has a number of very similar assets which must be replaced on an ongoing and rotating basis), past cost out-turns may be a very good proxy for the likely future cost out-turn.

In contrast, in the case of capex projects which are repeated very infrequently or not at all, the past cost out-turn(s) provide little information about the cost of capex projects in the future. A typical case of infrequently repeated projects are those reflecting the investments cycles stemming from the history of past network (under)developments. Under such circumstances, some (not necessarily all) companies may face a 'cliff-edge' investment problem due to a number of ageing assets reaching the end of their useful lives more or less simultaneously.

The regulator can undertake the following measures to mitigate this problem:

- measure capex via an annual 'capital consumption' based on annuities in order to smooth out the time profile of the capital consumption itself, and eliminate the implications of investments cycles;
- average out the annual capital consumption over a number of years;
- group the benchmarked companies in sub-samples with a similar asset age structure;
- apply a capital cost standardisation approach if there are differences in depreciation policies.

The non-controllable cost items stemming from externally imposed requirements will need to be removed from the totex cost base used in the benchmarking exercise. Regulators can decide to limit the backward-looking period to account for constrained controllability resulting from historical pre-conditions. This limitation, however, can induce disputes as to whether the threshold date allows for the fair and equal treatment of all the benchmarked companies.

Once total cost analysis is completed and revenue requirements have to be set, the gap between efficient allowed total cost (following benchmarking analysis) and capital adjustments must be bridged by analysing the need for top-up investment as determined by specific company-based and regional capex drivers.

Finally, quality regulation must be set up to counteract possible quality degradation incentives in response to the totex (or even yardstick) scheme. The interface of total benchmarking results with revenue requirement schemes, and the general logic of the price control should be kept consistent. This has been explicitly recognised by both the Dutch and Norwegian regulators. In these two countries, incentive schemes based on optimal quality of service targets have now been put in place.

6. REFERENCES

- AJODHIA, V.S., PETROV, K. and SCARSI, G.C. (2003). Benchmarking and its applications, *Zeitschrift für Energiewirtschaft* 27 No. 4, pp. 261-74.
- AJODHIA, V.S. and PETROV, K. (2004). Establishment of Cap Regulation and Determination of the efficiency factor 'X' for electricity networks. *European Transactions on Electrical Power (ETEP)* 14, No. 2, pp. 97-109.
- AJODHIA, V.S., PETROV, K. and SCARSI, G.C. (2004). Quality Regulation and Benchmarking. An application to electricity distribution networks. *Zeitschrift für Energiewirtschaft* 28, No. (2), pp. 107-120.
- BANKER, R.D., CHARNES, A. and W.W. COOPER (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science* 30, pp. 1078-92.
- BJØRNDAL, M. and K. JØRNSTEN (2003). Revenue Cap Regulation in a Deregulated Electricity Market – Effects on a Grid Company. *Central European Journal of Operations Research*, June 2003, 11, no. 2, pp. 197-215.
- CHARNES, A., COOPER, W.W. and E. RHODES (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research* 2, pp. 429-44.
- DTe (2002). "Maatstafconcurrentie, Regionale Netbedrijven Elektriciteit", DTe, Den Haag. Available in Dutch at www.dte.nl
- JAMASB, T., NILLESEN, P. and M. POLLITT (2003). Gaming the Regulator: A Survey. *The Electricity Journal* 16, pp. 68-80.
- JAMASB, T., NILLESEN, P. and M. POLLITT (2004). Strategic behaviour under regulatory benchmarking. *Energy Economics* 26, pp. 825-43.
- SEIFORD, L. and R. THRALL (1990). Recent developments in DEA: The mathematical programming approach to frontier analysis. *Journal of Econometrics* 46, pp. 7-38.
- SHLEIFER, A. (1985). A theory of yardstick competition. *Rand Journal of Economics* 16, pp. 314-27.
- WANGENSTEEN, I. and B.O. UTHUS (1999). Will income cap regulation and benchmarking affect the Norwegian distribution utilities' DA/DSM activities? DA/DSM Conference, Madrid, 28-30 September 1999.