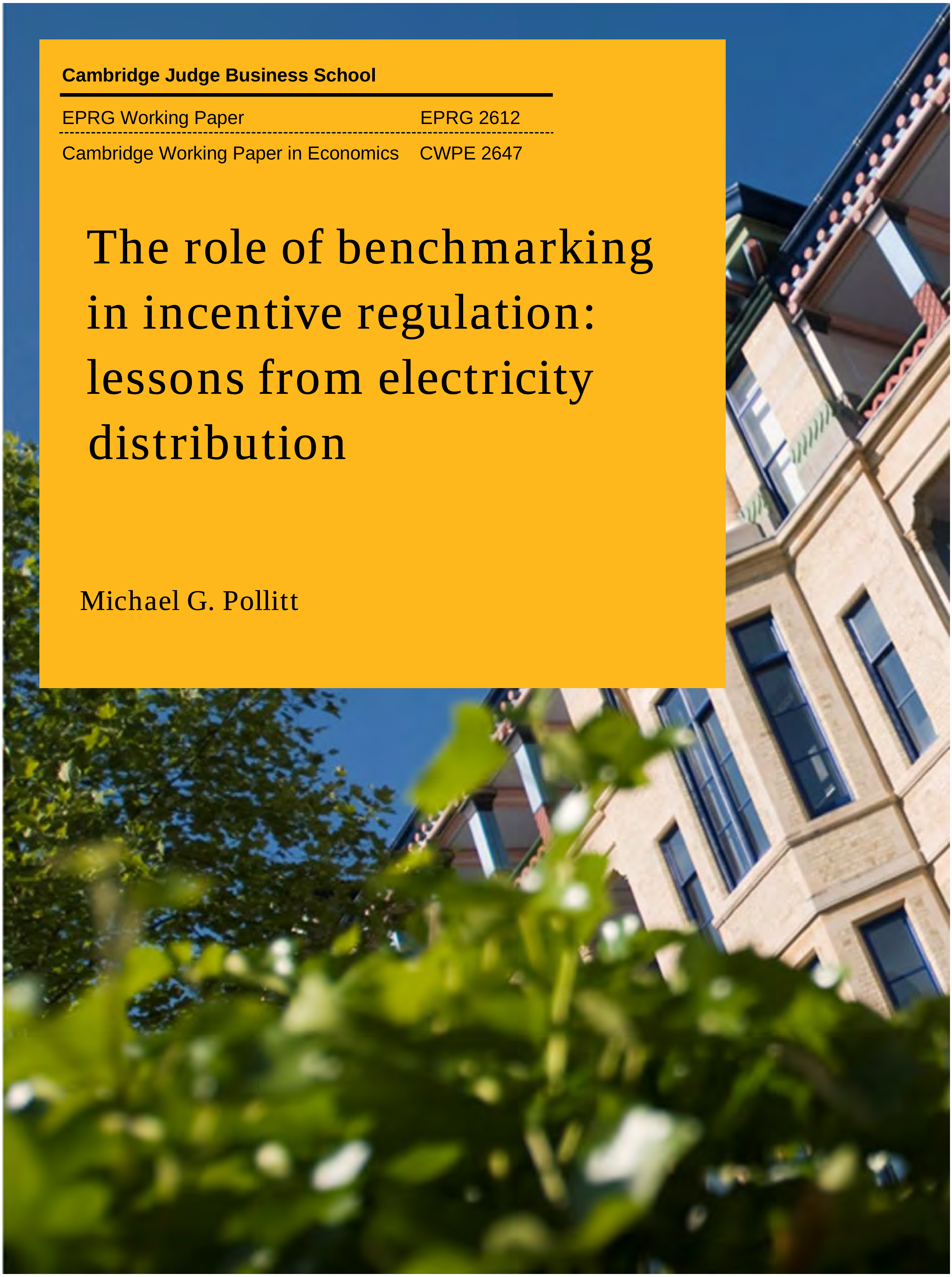
Cambridge Judge Business School

EPRG Working Paper EPRG 2612

Cambridge Working Paper in Economics CWPE 2647

The role of benchmarking in incentive regulation: lessons from electricity distribution

Michael G. Pollitt



Energy Policy Research Group

UNIVERSITY OF CAMBRIDGE
Judge Business School

Abstract :

This paper discusses the use of electricity distribution benchmarking within incentive regulation as part of price control reviews. It compares different methods, with an emphasis on efficiency benchmarking. We consider the choices facing the regulator in selecting a benchmarking method and in implementing it and how these may evolve over time. We contrast the experiences of regulators in Great Britain, the Netherlands and Germany in doing benchmarking within price control reviews and offer some lessons for the future. We suggest that benchmarking should be expected to evolve over time and that regulators should consider a 'benchmarking of benchmarking' with respect to their currently preferred benchmarking methods.

Keywords : DEA, SFA, COLS, frontier, efficiency, benchmarking

JEL Classification : L94

Contact	m.pollitt@jbs.cam.ac.uk
Publication	June 2026
Financial support	New Zealand Commerce Commission

Energy Policy Research Group - Cambridge Judge Business School

The Role of Benchmarking in Incentive Regulation:

Lessons from electricity distribution

Michael G. Pollitt[1]

Energy Policy Research Group

June 2026

1. *Introduction*

This paper discusses experiences with the use of benchmarking by electricity regulators to set revenue controls for electricity distribution utilities. Electricity distribution utilities are responsible for the lower voltages of the electricity grid and transport electricity from the high voltage transmission system directly to final customers. Kufeoglu et al. (2018) estimated there were over 7600 electricity distribution utilities in the world, in 175 countries. They are typically responsible for 15-25% of electricity system costs.

In 1993, Stephen Littlechild established a system for the benchmarking of electricity distribution utilities in Great Britain (GB) which continues to this day[2]. The GB has inspired a large number of other jurisdictions. This paper seeks to offer a reflection on the state of learning from the global experience of benchmarking electricity distribution utilities.

In Section 2 we begin with discussing the relationship between benchmarking and incentive regulation. In Section 3 we outline the options for comparative benchmarking. Section 4 compares efficiency benchmarking approaches. Section 5 discusses the choices faced by a regulator in implementing a chosen efficiency technique. Section 6 summarises some of the experiences in Great Britain, The Netherlands and Germany. Section 7 considers the future of benchmarking. Section 8 offers some conclusions for regulators.

[1] This version of the paper is part of the 2026 Festschrift for Stephen Littlechild. The author is grateful to the New Zealand Commerce Commission for funding the writing of an earlier version of this paper. He acknowledges the helpful comments of Diego Villalobos, Arik Mordoh, Bruce Mountain and his SCL AI agent and an anonymous referee. All errors are his own.

[2] See Littlechild (2024) for a discussion of the process of setting the first set of regulatory X factors for electricity distribution in GB. An early mention of the undertaking of both activity and regression based benchmarking is in Offer (1993, p.39).

2. Benchmarking and Incentive Regulation

Most incentive based regulatory systems for electricity distribution networks rely on CPI-X regulation as popularised by Stephen Littlechild (1983) in the UK and theoretically developed by Shleifer (1985).[3] It is important to point out that the key aspect of this is the fixing of the network revenue formula ex ante[4]. This creates an incentive to cut costs and improve service quality by allowing the regulated firm to keep profits for a period of 3-5 year (the time frame of a price control review). This has since been developed further by Ofgem, the energy regulator for Great Britain, into RIIO (Revenue = Incentives+Innovation+Outputs).[5] This emphasises the need for positive revenue incentives to support desired outputs (e.g. speed of connections) and innovation (e.g. via R+D projects).

It is important to acknowledge that a price control review can be viewed as a negotiation between the regulator and the regulated firm, where the firm is challenged to justify its stated revenue requirements following an initial suggestion from the regulator of what the allowed revenue should be (Beesley and Littlechild, 1989). Therefore, benchmarking can inform the initial negotiating position of the regulator. Regulatory benchmarking can be subsequently challenged by the regulated companies during the price control review process or by higher order regulators or the courts as part of the appeals against the price control review process.[6]

In this context efficiency benchmarking usually plays two key roles. First, benchmarking is about establishing starting points for target revenue at the beginning of a regulatory control period. Individual X factors can adjust for the different levels of potential for cost reduction based on how a regulated entity compares to an efficient version of itself (see Figure A.1). Second, dynamic benchmarking over several years can establish the trend level of

---

[3] As noted in Eyre and Pollitt (2016) a similar proposal was made in US regulatory context by Kendrick (1975) and Baumol (1982).

[4] The power of fixed prices to deliver incentives is noted in the context of US rate of return regulation by Joskow (1974).

[5] Ofgem (2010).

[6] In Great Britain, the Competition and Markets Authority (CMA) acts as the appeal body against the decisions of sector regulators on price controls (previously the Competition Commission, CC). The process for appeal is outlined in CMA (2022). For instance, Mid Kent Water appealed against price control on basis of non-use of panel data. CC (2000) ruled that it was difficult to decide between efficiency models.

productivity growth for a sector and set a general target level of efficiency improvement that can be pursued by all regulated entities, including the most relatively efficient firms.

Regulated companies can and do undertake their own joint benchmarking exercises, where they can share information confidentially about cost levels, without having to worry about regulatory oversight or the need to engage in regulatory gaming (making them more accurate and tailored to their commercial needs). Smaller companies might find a regulatory benchmarking attractive if it socialised the cost of undertaking an otherwise expensive benchmarking exercise.

Efficiency analysis on a comparator group of firms using multi-year data can give both an estimate of the trend productivity improvement and the relative inefficiency of a given firm. While there are a number of techniques that can provide estimates for both of these, it is not actually essential to do efficiency benchmarking to provide both individualised and general incentives to improve performance.

Performance-based rate making, as practiced in the US, might simply set a real revenue cap per customer with a revenue sharing mechanism (see Jamasb and Pollitt, 2001). The revenue sharing mechanism could adjust revenues each subsequent year in line a target rate of return, whereby if the rate of return increased the company would have to return some portion to the customers in lower prices. This would act as a rolling incentive to reduce costs and would encourage a more inefficient firm (which would find it relatively easy to increase profits) to reduce costs by more. This mechanism requires no external benchmarking and was, for example, implemented for Southern California Edison by the California Public Utilities Commission for 5 years from 1996, with annual X factors of 0-1.6% and four bands of revenue sharing around a target rate of return. For +/-50 basis points of rate of return, shareholder receives 100% of reward/risk; +/-50 bps to +/- 300 bps sharing between customer and shareholder of 25-100%; +/- 300 bps to +/-600 bps, 100% to shareholder; and above +/-600 bps the 'rates' are reopened by the regulator (see Jamasb and Pollitt, 2001). Joskow and Schmalensee (2025) suggest that Hawaii provides a good recent example of such a Multi-Year Rate Plan (MYRP), which has a similar structure of a 5 year X-factor with revenue sharing.[7]

Another simple mechanism that could be implemented as part of CPI-X regulation is an average productivity X factor. This was suggested by Uri (2002) for US telecoms. If the X

[7] The details of the plan are here: https://puc.hawaii.gov/wp-content/uploads/2020/12/PBR-Phase-2-DO-5-Page-Summary.Final_.12-22-2020.pdf

factor is specific to the industry being regulated and the industry starts off at a normal level of profitability, then in theory setting the X factor equal to the average productivity level will maintain reduce prices in line with costs and maintain profitability. One problem with this is that initially low profitability firms might never be able to hit their target level of profitability, as it does guarantee any individual firm will hit a target level of profitability. However, it has got strong incentive properties and can be set to run continuously. Setting a common X factor and adjusting on the basis of profitability (rather than comparative cost benchmarking) at the start of the regulatory period means relative inefficiency can persist (and grow) for long periods and is dynamically underwritten by upward revisions to prices to maintain profitability. Firms which are pursuing non-profit objectives (including a managerial quiet life) are underwritten by this system.

In practice many regulators have preferred to calculate individual firm efficiencies as part of their periodic reviews of revenue allowances for regulated electricity distribution firms (see Jamasb and Pollitt, 2001, 2007; CEER, 2026). A good reason for doing this is that this is what all firms do in setting internal performance targets. It has also been recognised that for smaller regulated firms, the regulator is providing information on potential comparators and performance that they might not otherwise have access to. It is also true that to some extent regulatory benchmarking can endogenously improve performance, by setting a target X to aim for, particularly in negotiations with unions or suppliers. The higher the X, ceteris paribus, the likely more cost reduction that can be induced from third parties working with the regulated firm. As Littlechild (2014) suggests, X is part of a 'rivalrous discovery process' where the regulator is pushing the firm to see how far it can reduce costs.

3. Comparative Benchmarking Options

A number of comparative benchmarking options can be pursued.

Activity benchmarking is a standard management consultancy style of benchmarking. This can be functional or task based. Functional is where IT departments or Head Quarters staff are compared directly across firms (who do not need to be in the same industry). Task (or process) based might look at the cost/time/labour to fix a fault in the network or install a new transformer. This sort of analysis can be undertaken by engineering or economic consultants on behalf of the regulator. This was being undertaken by the GB electricity regulator in the 1990s.[8]

[8] Noted in Jamasb and Pollitt (2001).

Matching analysis might compare key performance indicators (KPIs) across matched firms. This might look at the number of staff per customer for firms of the roughly the same size, with similar customer types (e.g. two utilities covering large metro areas). This is less popular among regulators who don't have national matched pairs, than it is among the companies themselves. Norway benchmarked its electricity transmission utility with the one in Sweden.[9] EURELECTRIC commissioned a study comparing KPIs for island electricity systems.[10]

Model analysis uses an engineering model to build an efficient network to serve the actual customers of a given network. This 'ideal' model or 'norm' model has been used in certain countries (e.g. Brazil[11] and Sweden[12]).

Frontier efficiency analysis (see Coelli et al., 2005; O'Donnell, 2018) has become popular across the world (see for example Jamasb and Pollitt, 2001, and Haney and Pollitt, 2009, 2013a). Three techniques have been popular: data envelopment analysis (DEA), corrected ordinary least squares (COLS) and stochastic frontier analysis (SFA). Historically we have seen a lot of use of DEA (though less than activity benchmarking); some COLS; and less SFA (see Table 1).

[9] Noted in Jamasb and Pollitt (2001).

[10] Cross et al. (2017).

[11] For instance in the Brazilian electricity regulator, ANEEL undertook a benchmarking exercise around 2004 for 61 electricity distribution utilities using an engineering model to assess what their costs should have been given their configuration (see, Silva, 2007).

[12] See Jamasb and Pollitt (2008) on Sweden.

Table 1: Use of different benchmarking techniques across 43 electricity regulators (Haney and Pollitt, 2009).

Overall, there is weak evidence of experience and neighbourhood effects in best practice in regulation (c.f. Haney and Pollitt, 2009, 2011). Thus, similar groups of countries have a tendency to use similar methods in their benchmarking. We discuss why countries have particular methods over certain periods below.

Multi-factor Total Factor Productivity Analysis (MTFP) can also be used to calculate relative productivity and trend productivity growth. This is involves fixing the input and output weights to calculate weighted productivity score and annual rate of change of this score. Such an approach is similar in spirit to DEA with no allowance for stochastic factors, but with fixed weights on inputs and outputs. This is one of the methods used in Australia to cross check other methods (Frontier Economics, 2023).

For an up-to-date summary of regulatory approaches to benchmarking among EU-27 energy regulators, see CEER (2026). Frontier Economics (2023) discusses the details of recent benchmarking experiences in GB, Finland, Australia and Germany.

In addition to the above methods for looking at the efficiency of individual firms, we see other approaches being used to look at productivity trends in setting industry level X factors. For instance, the regulated water industry in the UK has had its costs allocated across other industrial sectors to construct a benchmark real price effect index which is a weighted average of non-water sectors (see CEPA, 2024a). It is also possible to set the target cost level for each firm on the basis of average industry costs (this currently happens in the Netherlands). This is quite close to the yardstick scheme originally envisaged by Shleifer (1985) and involves no direct comparative efficiency calculations.

When resources are not a problem or comparators are difficult to find domestically non-frontier efficiency methods have a lot of value, because they can deal with small samples of comparable domestic regulated firms.

When comparators are large enough in number DEA, COLS or SFA are implementable. Individual regulators normally favour using domestic samples for these efficiency studies, though they periodically may look to international comparisons to check how far from the frontier their domestic firms are. This is because justifying setting price adjustments or X factors on the basis of international comparisons is difficult. A more reliable route to gaining best practice is allowing foreign takeovers or mergers. For instance, National Grid in the UK did takeover US distribution firms, while ENEL have taken over a number of distribution utilities around the world.[13]

4. Comparing efficiency benchmarking approaches

DEA has the advantage of letting the data to directly determine the frontier of efficiency and of allowing comparators who are defining the relevant portion of the frontier to which an individual firm is being compared to be identified. DEA is essentially an extension of single factor productivity measures and simple comparisons of unit costs that might be done by companies, themselves. Efficiency can be visualised and it is the method which is closest to what firms do when internally comparing their performance to their peers.[14] It does not require a functional form to be imposed on the production process and can incorporate variable returns to scale (or impose constant returns to scale[15]). It is also easy to incorporate multiple outputs and inputs and to look at the evolution of productivity over time. Productivity can be further decomposed into frontier shift caused by the leading firms and catch up of an individual firm moving closer to the frontier. The advantage of a DEA productivity decomposition being that no assumptions are required about the pattern of productivity growth over time. Adjustments for environmental conditions can be incorporated directly into the DEA or adjusted for in secondary analysis. DEA has the disadvantage that efficiency scores are susceptible to outliers. Therefore, it might exaggerate efficiency scores or productivity growth by allowing frontiers to be defined by

[13] https://www.enel.com/company/about-us/where-we-are

[14] I was once involved in an exercise where we used DEA to identify one particular firm which could serve as a benchmark for one of the firms we were analysing (on behalf of the holding company CEO). Having initially claimed that there were no relevant peer firms, the relevant subsidiary manager admitted that they were already using the firm we had identified using DEA as a comparator.

[15] Constant returns to scale assumes that optimal subdivision is possible at higher than optimal scale.

data which might reflect one-off reductions in costs or increases in output with no adjustment for the possibility of error. Frontier stripping where the entire group of frontier firms is removed and scores recalculated, can be practiced to remove outliers and test the stability of the frontier.

SFA takes explicit account measurement error in the data. A functional form can be defined (usually Cobb-Douglas or Translog) on the production function and productivity growth can be identified, though some assumptions are required about the nature of the trend (e.g. whether it is a single value over the period, or rising or falling). SFA does tend to estimate higher efficiency scores as it is likely to attribute deviations from the frontier to measurement error rather than efficiency. Efficiency scores tend to be higher than under DEA. The more complex the estimated functional form of the cost function, the more likely high efficiency scores are to result. It is worth pointing out that parametric approaches inherently try to maximise fit and hence 'despise' error. SFA can identify time trends in panel data in order to identify productivity growth rates (e.g. Greene, 2005). Thus we can say that while DEA would likely exaggerate efficiency scores and frontier productivity growth, SFA would likely underestimate them.

COLS (or MOLS)[16] is another parametric approach which is practiced by regulators, which fits a functional form to the cost function but shifts it to incorporate all of the firms. It can be more or less generous in terms of inefficiency, depending the extent to which the shifted cost function envelops all or only, say, 75% of the less efficient data, leaving the 25% to be deemed efficient. This approach has been used extensively by the GB regulator, Ofgem. COLS tends to be used to calculate efficiency, rather than productivity.

The Finnish regulator has used the Stochastic Non-parametric Envelopment of Data (StoNED) approach (see Johnson and Kuosmanen, 2015). This is an approach which estimates a non-parametric fronter using minimal axiomatic assumptions about the functional form of the production function. It is an SFA type approach that attempts to combine the advantages of DEA (which does not impose a functional form) with the advantages of SFA (in allowing for measurement error). Andor and Hesse (2014) conduct a simulated comparison - across 9400 datasets - of this approach with DEA and SFA. They conclude it does perform relatively well especially when there is 'noise' in the data, though SFA performs generally well. They also find that DEA performs relatively well when there is no noise in the data.

[16] Modified Ordinary Least Squares. See Parmeter (2021) for a fun discussion on the origins of the terms COLS and MOLS, finishing with a call to just use the term COLS.

All frontier efficiency measures can be adapted to incorporate multiple variables which are deemed to be cost drivers (these can be incorporated as inputs, outputs or environmental variables depending on the nature of how they affect costs). However, in general more variables means higher efficiency scores and also the risk of spurious correlation. As we discuss below which variables should be included is an important area of disagreement between regulators and their regulated companies (and among the companies).

Many regulators have used a variety of approaches in order to try and identify efficiency and productivity. These can include DEA, COLS and SFA or some combination of an efficiency approach and other benchmarking approaches as detailed in the previous section.

Lovell (2006) discusses best practices in efficiency analysis in theory and practice. He points out that other advanced variants of efficiency exist, such as Stochastic Data Envelopment Analysis (SDEA)[17] or versions of both DEA and SFA which incorporate environmental variables. Best practice in implementation involves: a large / high quality dataset; panel data; consistency with engineering / well behaved functional forms; use of bootstrapping / confidence interval analysis; checking for consistency with non-frontier methods; inclusion of quality / environmental / input price variables; and clear value added in doing efficiency analysis. As Weyman-Jones (2006) observes about the use of efficiency analysis in practice: 'The overwhelming impression of regulatory and governance case studies is that sample size, variable choice, model specification and choice of methodology has been governed by different objectives from those in the theoretical literature.' (p.25). Thus actual benchmarking is a lot messier in practice than would be desirable in theory.

5. *Regulatory choices in implementing a benchmarking method*

The peer group

One key choice is which group of firms will be benchmarked against which. For a country with a large size variation of firms a decision might be taken to drop the smallest and largest firms from the sample. This is what would happen in most analyses of large samples of data, where the concern was to accurately estimate the best fitting statistical relationships, rather than to include all of the data points. Indeed it might be possible to partition the data into sub-groups and estimate DEA or SFA on each sub-group. Statistical techniques exist to allocate firms to groups (rather than arbitrarily using one of the parameters, such as number

[17] This adjusts a DEA frontier to allow for stochastic error in the underlying data. See, Weyman-Jones (2001) who discusses the application of Olesen and Petersen (1995).

of customers). Thus, Latent Class analysis which sorts the sample into a number of groups should these be significantly different (see Llorca et al., 2014).

The outputs, input and environmental variables used

Electricity distribution utilities deliver electricity to a number of customers over a specific area at a given quality of service. They use physical assets, labour and materials do to this. Costs can be measured in real money or in physical units. Customer mixes (i.e. the ratio and size distribution of business customers), degree of undergrounding of wires, peak demand, peak capacity, geography and weather may all be important drivers of cost. It is also often preferable to work with a smaller number of variables on the initial analysis (because the larger the number of variables the larger the efficiency scores, due to the lower degrees of freedom). Some variables may be highly correlated and hence do not all need to be included. More variables also increases the chance that estimated cost functions are badly behaved (costs appear to fall in variables which are clearly cost increasing). Starting with a small number of variables makes it possible to test whether other variables are significant and make a second-stage adjustment or to exclude them entirely due to insignificance. Including more variables directly in the efficiency analysis biases measured efficiency up, due to effectively reducing the number of firms against which any given firm is being compared. While preferred measures of output tend to favour measures of outputs directly demanded by consumers (i.e. energy and connection), it is possible to use measures of capacity (e.g. total network length) as outputs, to reflect that the network is built ahead of need. One good regulatory strategy is to ask companies which variables they think are materially significant and then test whether there is evidence that this is the case.

How to incorporate individual variables

Efficiency analysis has to decide how to incorporate relevant variables. Variables can enter efficiency analysis as main outputs or inputs or as secondary variables. This is particularly true of environmental or quality variables which could be treated as different types of variables. For instance, DEA customer minutes lost (CML) can be treated as an input or as an environmental variable within the analysis. The particular mode of inclusion may have implications for the adjustments that need to be made to the variable prior to inclusion. For instance, DEA scores should be scale invariant and this means that if CML per customer should not be used, and such a variable needs to be scaled by the actual number of customers.[18]

Which costs are comparable

[18] See Yu et al. (2009).

An important up-front choice is to what cost data exactly is benchmarking applied. Is it the entirety of the costs or some fraction of them? Is it physical inputs and outputs rather than financial measures of cost, to maintain comparability and avoid accounting problems? When it comes to benchmarking costs, it is usual to focus the benchmarking on operating costs (opex) rather than capital costs (capex), but it is possible to benchmark total costs (totex = opex + capex), or to even benchmark opex + rental cost of capital (which would be driven by the historic fixed assets).

If benchmarking is only applied to opex, this does raises issues of opex-capex trade-offs and still leaves the question of how to ensure capital investment is efficient open. GB addressed this with menu regulation of capex expenditure, which later became menu regulation of totex expenditure. Benchmarking is useful for determining the baseline revenue against which rewards and penalties under menu regulation are determined. Meletiou et al. (2023) find that 63% of the 77% of EU-27 countries applying efficiency benchmarking applied it to opex and 37% to totex.

Cost adjustments to improve comparability

There is also the issue of the extent to which pass-through or firm specific costs are removed or adjusted for in the data, pre-analysis. Ofgem do multiple adjustments to their data in order to make the data comparable before doing the benchmarking analysis, and then have to apply measured efficiency score to that portion of the data which was actually used in the efficiency analysis.

In its review for the 2005-10 electricity distribution control review Ofgem had 14 regulated entities (7 independent), used COLS (1 output/1 input), did not use panel data or confidence interval analysis. It did cross check with other methods (DEA, SFA), adjusted the COLS to the upper quartile of regulated firms and made 13 adjustments to the data they received from the companies before undertaking the efficiency analysis.[19]

If costs are benchmarked, a key issue is the extent to which local variations in labour costs are taken into account assuming a common cost of capital and materials costs. It would be usual to make an adjustment for major city costs if wages there are significantly higher than the national average. Though there remains a good argument for not doing this as higher wages may reflect higher labour productivity and hence outturn unit labour costs might be expected to be the same across a country.

[19] See Pollitt (2005).

As an example of the overall revenue impact of benchmarking. Here is an example from the GB electricity distribution price control from 2005-06. Operating costs are benchmarked using COLS (See Pollitt, 2005).

Regulated entity: United Utilities

£67.1m (2002-03): Normalised actual costs in benchmarked year.

£54.8m (82%): Efficient costs

£205.2m: Allowed revenue in 2004-05, last year of current price control period

£67.0m: Allowed operating costs in 2005-06 (including non-comparable opex with difference being mainly local property taxes)

£220.9m: Allowed revenue in 2005-06, first year of new price control period

Thus only around 33% of actual revenue in 2002-03 was frontier efficiency benchmarked[20]. For comparison here are some different sources of regulated revenue impact:

+/- 5% on efficiency score:	+/-	£3.4m
+/- 1% p.a. on frontier shift:	-/+	£2.7m (by year 5)
+/- 1% on rate of return:	+/-	£9.2m
+/- 10% on capex:	+/-	£11.2m

If multiple efficiency scores for a single firm are calculated the issue is the extent to which they are adjusted by averaging or taking the highest efficiency score (usually the lowest would be considered an outlier). In the German example below, the highest score of the four scores calculated was used.

Translating efficiency scores into X-factors

There then remains the issue of how to translate a final calculated efficiency score of, say, 80% into the X factor within the revenue cap formula. As Ros et al. (2024) point out there is little theoretical and empirical support for mechanically translating backward looking efficiency scores directly into forward-looking X factors.

[20] Assuming that the revenue in 2002-03 was similar to 2004-05, during same price control period.

One extreme approach would be to adjust the revenue from the first year of new price control period down by an amount consistent with a move to 100% efficiency. This would be a maximal P0 (initial price) adjustment. X could then be set equal to the expected productivity growth of the most efficient firms (which might be of the order of 1% p.a. in an advanced economy).

Another approach would be to determine an acceptable final expected inefficiency gap by the end of the price control. Say this was to close 50% of the gap by year 5 of a 5-year price control period. Assume the gap was closed linearly. This might result in an X factor equivalent to improving efficiency by 2% in each of the 5 years to close the gap plus a general productivity growth factor of 1%.

These two paths for X are very different in terms of end points for revenue and the pathway of revenue during the 5 year price control (see Figure A.1). The second would involve substantially higher revenue over the price control period for the regulated firm. Both represent different degrees of confidence in the ability of a regulated firm to cut costs over a five year period on the basis of a regulatory assessment of efficiency.

Other approaches are possible. Following the initial price control of distribution companies in the Netherlands in 2000, the regulator (DTe) directly translated efficiency scores into revenue with a simple cap of 8% on X (see below).

An obvious mistake to avoid is misapplying efficiency scores to investments which should not be part of the benchmarking exercise. There has been an issue for both Netherlands and German utilities that the regulator calculates efficiency scores based on expenses occurred in advance of the outputs materialising. Thus if costs are incurred in advance of need this could lower an efficiency score and have the effect of appropriating the return to an investment. By focussing efficiency analysis on opex (as in GB) this is much less likely to occur.

Regulators have used benchmarking to establish productivity growth in the sector. For instance in 2018 Ofgem undertook an exercise to look at long run productivity trends in the electricity sector using DEA (from 1990)[21]. These exercises resulted in the identification of a long run trend of around 1% p.a. in total factor productivity. This number then appeared as

[21] Ajayi et al. (2018).

an efficiency target for all firms in the subsequent price control[22]. This work also found that including a wider range of quality and environmental outputs improves measured productivity growth in a sector which has been encouraged to improve quality of service and environmental impacts.

6. *Lessons from leading jurisdictions on their use of benchmarking of electricity distribution*

Great Britain

Great Britain (GB) has made use of COLS based benchmarking of its 14 distribution utilities. This was very successful initially because the privatised companies were initially very inefficient (see Ajayi et al., 2022).

Ofgem (and its predecessor regulator, Offer) has put a lot of effort into data cleaning and adjustment to make comparisons of the operating efficiency between the firms. There was some attempt to ensure that 100% efficiency was not expected to be achieved initially. For the price control from 2000-2005, firms only had to remove 75% of their inefficiency by the second year. As we saw from the United Utilities example above, benchmarking was only one regulatory impact on revenue. Incentives for quality outperformance were also significant. Ofgem's extensive engagement process with individual firms has meant that there has been a high degree of trust in the process especially with respect to data cleaning. However. the significance of benchmarking exercises for individual company review targets has declined over successive price control reviews.

Ofgem has evolved its use of benchmarking towards a pooled OLS approach making use of past and future projected data. This retains an emphasis on a straightforward regression but now incorporates more years of data with a time-trend (see Ofgem, 2022b, c)[23]. A good summary of how it used these approaches within the overall setting of X factors in the current price control (2023-2028) is provided by Oxera (2022).[24]

For the latest price control period 2023-28, Ofgem made use of four models – three based on a composite scale variable (CSV) COLS approach and one based on disaggregated activity benchmarking. The included actual data from 2016-21 (5 years) and simulated data from

[22] Ofgem (2022a), *RIIO-ED2 Final Determinations*, Nov. 2022, p.12.

[23] For a discussion of the latest methodology see Ofgem (2022b, p.246ff).

[24] https://www.oxera.com/insights/agenda/articles/riio-ed2-final-determinations/

2022-28 (6 years). Two of the CSV models use actual plus simulated data, the third only uses the simulated data. The weighting placed on the results was 16.67% on the first three and 50% on the latter. The simulated data incorporates expected take-up rates of EVs and heat pumps. In translating these into X factors, the assumption is that each of the 6 regulated firms will reach the 85% percentile of efficiency by year 3 of the price control period. The use of simulated data suggests that role of actual cost performance in setting the X factor is much diminished. The large weight placed on the activity benchmarking shows a reduced role for whole company cost benchmarking.

There is a general assumption of a 1% p.a. trend improvement in efficiency for 2023-28.

The Netherlands

In the Netherlands, the energy regulator (Dte) undertook efficiency analysis of electricity distribution companies for the first time in autumn 2000 (see Nillesen and Pollitt, 2007). A July 1999 consultation document lays out ambitious plan to benchmark transmission and distribution utilities by Dte (the Dutch energy regulator).

The regulator analysed 19 electricity distribution companies using DEA, with no panel data or benchmarking. They did not cross-check with other methods and made no general adjustments for environmental factors. They capped the 3 year X-factor at 8% p.a., however it was assumed that for companies who were not capped that they had to remove 100% of inefficiency by the end of the three year price control. Legal challenges were launched mid-2001 with appeal to the NMa (Dutch competition authority) against DTe by companies. DTe decisions overturned by the NMa and revised in 2002. Legal challenges led to 300 court cases! This reaction was surprising, especially as the network companies have remained in public ownership. It can be attributed to a poor relationship with companies due to the rapid timetable, an initial lack of education about the process and poor regulatory public relations. The regulatory strategy of DTe was not incentive compatible as it did not focus on revealing the scope for long-term cost reduction and was excessively short-termist. The Dutch regulator (now the ACM) has not used frontier efficiency benchmarking for individual electricity distribution company X factors in subsequent price control reviews, but moved to a yardstick approach.

The yardstick approach involves the calculation of a composite measure of output for each firm and for the industry as a whole (see Montfoort et al., 2024). The current composite output index includes measures of connections, transportation service and feed-in service (e.g. distributed generation). The first two include 40 different measures. The benchmark composite unit revenue is calculated on basis of the previous three years. This provides the baseline regulated revenue per composite unit. The totex distribution costs, excluding pass-through transmission costs are then used to calculate a unit cost. The fact that companies are collecting more revenue than they need to cover current industry costs is one source of a positive X factor. The other is the expected frontier shift in the industry. The ACM has looked at long-run productivity trends in comparator industries to identify long run dynamic efficiency, which is very low at 0.1% for the period 2000-2017[25]. There have been 8 regulatory periods to date, with the first 6 being for three years and the latter 2 being for five years. The current period is 2022-26.

The performance of the Dutch electricity distribution sector over the period 2000-24 is reviewed in Nillesen and Pollitt (2026). They find that after initially strong performance in terms of cost reduction, average costs have been rising since 2017. They show that the early headline reductions in average costs were largely driven by falls in the regulated rate of return. Thus rising interest rates and investment requirements will overwhelm any reductions arising from yardstick incentive regulation and will not prevent rising residential electricity network bills in the next regulatory period.

Germany

The German energy regulator, the Bundesnetzagentur, undertook its first efficiency benchmarking in 2007.

They initially analysed 328 electricity (and 488 gas) distribution companies using DEA, SFA and COLS for a single base year. After the elimination of outliers the final useable sample was 205 electricity distribution utilities. They did do bootstrapping to get more robust estimates and did include environmental factors. In using the results of the estimated scores (they used the highest of the scores from each of four methods), they capped the maximum efficiency improvement over 10 years at 50%. This process was initially overseen by two leading academics of efficiency analysis: Per Agrell and Peter Bogetoft[26].

In their review of the first six years of the German experience, Agrell and Bogetoft (2014) note the importance of 'cautiousness' which they attribute to the choice of the highest score from each of the four methods with a minimum allowed efficiency of 60%. The four methods used for comparison were a DEA (assuming CRS) and an SFA approach each calculated with using two different measures of Totex (opex + capex). The different measures of included either the nominal value of capex or a standardised measure. Companies are expected to eliminate all of their measured inefficiency by the end of the price control period.

[25] See Cunningham et al. (2020, p.v).

[26] See Agrell and Bogetoft (2007).

This approach has been implemented over four price control periods running from 2009, with the current period 2024-2028[27]. This continues to exploit the large dataset that is available to the German regulator.[28] An interesting feature of the German system is that smaller distribution companies can opt out of the benchmarking exercise and be given the average efficiency score of the included companies, under a simplified procedure.[29]

The regulator has proposed a new system of regulation from 2029 which would reduce the significance of benchmarking (but not eliminate it) under a new system of regulation know as NEST – 'Networks – Efficiency – Security – Transforming'.[30] Among the changes proposed are the dropping of the use of DEA, in favour of only using SFA scores from the two different measures of capex, and the raising of the minimum efficiency score to 70%.

This follows a 2021 decision of the European Court of Justice that ruled that the regulator was not sufficiently independent from government in its ability to set regulated tariffs (inter alia), because of the overly prescriptive nature of the legislation covering its activities.[31] Indeed the relevant law, the ARegV, Annex 3[32], very unusually mandates the use of DEA and SFA and details how they may be used, rather than granting regulatory discretion on how revenue caps are calculated.

Regulated firms are subject to a general productivity factor of 0.91% p.a. in the current period which has been set using Malmquist and Tornquist analysis of historic data.[33]

---

[27] See Trinker (2004).

[28] A short summary of the current approach, from the regulator, is here: https://www.bundesnetzagentur.de/EN/RulingChambers/Chamber8/RC8_05_Revenue%20caps_revenue%20regulation/54_Efficiency%20benchmarking/Efficiency%20benchmarking_.html

[29] https://www.bundesnetzagentur.de/EN/Areas/Energy/GeneralInformationRegulation/IncentiveRegulation/Tools/start.html

[30] See: https://www.bundesnetzagentur.de/SharedDocs/Pressemitteilungen/EN/2025/20251210_NEST.html

[31] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:62018CJ0718

[32] https://www.gesetze-im-internet.de/aregv/anlage_3.html

[33] https://www.energate-messenger.com/news/246693/xgen-to-remain-almost-stable#:~:text=Bonn%20(energate)%20%2D%20The%20Federal,Index%20and%20the%20Törnqvist%20Index.

General observations about the practice of benchmarking electricity distribution companies

There is a role for comparison of the results of benchmarking methods, especially to justify why a given method is representative or fair to the companies. Bauer et al. (1998) argue for observable consistency between approaches, namely: different approaches should have comparable means, standard deviations and distributional properties and rank orders, with the same firms at the top and bottom; demonstrable stability over time; consistency with what the market thinks (i.e. if the market thinks a firm is well-run we might question consistently poor frontier efficiency scores and vice versa); and that frontier method results should be consistent with results from non-frontier methods.

Some countries, such as GB, are doing a decent job with the data that they have and simpler methods may serve a useful purpose within the context of a trusted negotiation process around the price control.

There is an issue with the failure of more sophisticated methods to deliver useable results on small samples of data. While Lovell (2006) is right about what makes for a more robust efficiency analysis, in practice getting data of this quality is difficult and even then it may not be useable with respect to all individual company efficiency scores.

In practice most efficiency exercises are naturally generous to companies because adjustments are almost always in favour of company revenue and because benchmarking exercises don't include all the comparator companies in the world (e.g. the utility in London does compare itself to those in New York and Paris, but the regulator does not).

There is (and should be) evidence that scores are converging over time[34] and hence firm specific efficiency scores are becoming less material to revenue.

The percentage of allowed revenue subject to benchmarking is likely to fall as the value of comparable operating costs falls relative to depreciation and returns on regulatory asset bases.

[34] German average electricity distribution efficiency scores improve from 92.2% in 2008 to 94.7% in 2013 to 94.1% in 2019 and then 95.9% in 2024 (see Agrell and Bogetoft, 2014, and Trinkner, 2024).

There continues to be a lack of use international data in the regulation of electricity distribution.[35]

As we observed in both the Netherlands and Germany, benchmarking processes are subject to ongoing legal obstacles and challenges. The imposition of X factors in Iceland – informed by DEA - were successfully appealed against for the most recent period (CEER, 2026)[36]. In the US it is generally difficult for a given regulated firm to be regulated on the basis of another firm's costs (hence the use of PBR arrangements) and comparative benchmarking is not used (see Joskow and Schmalensee, 2025)[37].

7. *The future of benchmarking*

A concern of those who do benchmarking is whether repeated benchmarking is becoming harder to justify over time, especially when it comes to using individual firm efficiency scores.[38] This is in addition to those (e.g. Cronin and Motluk, 2007) who think that comparative efficiency benchmarking does not accurately mimic the pressure of market to improve performance in the same way that a single yardstick does in Shleifer (1985).

This is because the underlying cost function becomes more difficult to model as both the outputs and inputs of the firms proliferate and differentiate. This is especially true with respect to distributed energy resources and innovation. Distributed energy resources (DERs) can vary significantly – in quantity and type - from network to network (e.g. amount of batteries, PV, small wind, smart demand side response) making it difficult to assess their impact on efficient costs. Some distribution networks (particularly larger ones with favourable geography) may have substantially more opportunity to use market based procurement of flexibility to optimise their networks, this may result in them looking very efficient in terms of the provision of their core service, when actually they have outsourced some service provision – reducing regulated opex and capex – to the 'market'. Innovation

[35] See Agrell and Bogetoft (2018, Table 16.1), who suggest that among reporting European countries using frontier efficiency techniques, only Iceland uses international benchmarking of its DSOs. In CEER (2026) only Portugal (among 24 European country members in Table 2 below) reports that it used an international sample in the calculation of frontier efficiency scores.

[36] The basis of the appeal was the lack of a legal basis for the imposition of X factors (CEER, 2026, p.88).

[37] This is because US legislation and case law protects regulated firms from regulation that threatens a firm's financial viability (see Weisman, 2019). This in practice means that firm specific X-factors need the buy-in of regulated firms. I am grateful to Dennis Weisman for a discussion on this point.

[38] See for example Agrell et al. (2013).

projects may be funded by specific innovation grants and have the effect of reducing core network costs, however benchmarking may not be capable of distinguishing between companies that have received innovation funding (e.g. to put in an Automated Network Management (ANM) system to better optimise the use of existing assets at no impact on costs) and those that have not. Similarly, if increased stakeholder engagement suggests some electricity distribution company customers are willing to pay for more undergrounding of wires this may increase costs and measured inefficiency under conditions where this reflects customer preferences. The above are examples where firms are doing fundamentally different things on the basis of differing opportunity and local preferences and cannot be regarded as having access to the same technological choices to be efficient (which is the fundamental assumption of all comparative benchmarking). This would also be an issue for a common total factor productivity (TFP) based X factor, where an electricity distribution utility in an area with a relatively high preference for undergrounding would have lower TFP growth than the average.

This sort of increasing diversity – which itself reflects different circumstances between utility service areas - means that whole company comparisons of efficiency are difficult to justify without more serious attempts to focus in on core service costs and make relevant adjustments for favourable innovation funding or DER environments[39]. This argues for more bespoke targets for efficiency based on the specific circumstances of the company rather than those based on a common concept of the efficient firm. While DEA and SFA remain of value in testing for general impacts (e.g. do innovative firms have higher efficiency?) and in estimating productivity trends, individual efficiency scores are likely more and more open to stakeholder and legal challenge.

Ofgem's RIIO process downgraded the role of benchmarking by suggesting that negotiated regulatory settlements, procurement auctions, private wires networks and innovation were alternatives to a regulator led process of cost assessment (see Pollitt 2010). More recently, Ofgem has been considering the future of the price control process once again from 2026, with increased emphasis being put on the need for regulatory learning across the price control period, including with respect to uncertainty mechanisms and adaptive regulation. Indeed, Pollitt et al. (2024) emphasise the need for a 'learning regulator' who has the capacity to use past information (as with conventional benchmarking), learn from stakeholders in real time (about issues arising within price control periods) and be prepared for future learning points in the mid-term (e.g. when large context dependent investments, such as support for EVs or Heat Pumps, might need to be initiated).

[39] Europe Economics (2023) makes a similar point in the context of a discussion about the future of network regulation in Denmark.

Advanced country electricity distribution networks are struggling with rising costs and falling productivity (see: Ajayi et al., 2022, on Great Britain; CEPA, 2024b, on New Zealand; and AER, 2024, on Australia). Headline charges are further increased by the rising cost of capital which has started rising after many years of falling. A good example being the Netherlands where the real weighted average cost of capital (WACC) has had to be increased from a low of 1% in the most recent price control.[40]

Any regulatory benchmarking method is subject to a Lucas critique[41], as with all policy based on previous econometric (or estimated) relationships, i.e. it is good to begin with but then it is likely to become less effective over time. It is therefore a good idea to try it for a period and then drop for a period and then revisit. Prior to 1990 frontier efficiency analysis was mostly used for hypothesis testing, rather than firm specific efficiency scores, increasingly we likely to be going back to the former use.

It is interesting to end this section by comparing the use of frontier efficiency techniques over time (see Table 2). We start with the 24 European countries in Haney and Pollitt (2009) and compare these with more recent data for 2025 (based on CEER, 2026). Two observations can be made. First, the use of frontier techniques is not expanding. 10 countries were using such techniques in 2009, but only 8 in 2025. Second, both DEA and SFA have declined in popularity, to be partially replaced with StoNED and more COLS/MOLS.

This table and our earlier case studies suggest that historic benchmarking techniques are thought to have continuing use value by some European regulators but that such a positive view its use value is not shared by all regulators.

[40] https://think.ing.com/articles/dutch-utilities-addressing-the-capex-and-opex-challenges/

[41] Lucas (1976).

Table 2: Use of benchmarking techniques in regulation of European electricity DSOs: compared between 2009 and 2025

Y = in use in latest regulated revenue setting.

(Y) = indicates use, followed by subsequent successful legal challenge.

Sources: Haney and Pollitt (2009), Agrell and Bogetoft (2018), Europe Economics (2023), Rita et al. (2023), CEER (2026).

8. Conclusions for regulators

Electricity distribution firms benchmark themselves (or should do), therefore there would need to be a good reason for a regulator not to periodically check the relative efficiency of its regulated firms, against each other and international benchmarks. Any regulator which is not undertaking comparative efficiency analysis (e.g. New Zealand) of its electricity distribution utilities should do so. A one-off benchmarking exercise for such a regulator might be useful in establishing whether there are currently wide divergences in performance and hence whether future use of comparative benchmarking might be useful.

As in Great Britain in 1990, the likelihood is that initially there are substantial gains from implementing comparative benchmarks to reduce costs.

Regulators can undertake all sorts of types of benchmarking simply for information, particularly benchmarking exercises involving international samples of firms[42]. However much greater care must be taken when a benchmarking analysis is going to be used to directly inform the setting of X factors. It is here that legal defensibility, based on best practice in practical implementation, is important.

A variety of benchmarking techniques exist, from activity benchmarking to frontier efficiency analysis. It is good practice to use several techniques to cross-reference the results. This is because this will enhance confidence in the efficiency targets set, either by suggesting that alternative methods give similar results or that there is a good reason why the methods relied on are more reliable (e.g. a finding that DEA or SFA gives implausible results and should be dropped). It is recommended that where multiple methods are used and considered equally reliable the highest efficiency score is used in X factor calculations.

A regulator can and should do reasonable data cleaning in order to ensure improved comparability between its regulated firms. This has the additional advantage of better understanding regulated firm costs and financial reporting. This is likely to drive firms to report on a similar basis making future comparisons easier.

Whichever frontier efficiency techniques for efficiency and productivity calculation are adopted, some allowance should be made for errors in measurement and representativeness of potential non-representativeness of benchmarked years. This can be done statistically within the chosen technique (e.g. via use of SFA, only setting the COLS frontier to the first quartile of efficiency or frontier stripping within DEA) or via adjustments ex post (e.g. setting a maximum value of X or assuming less than 100% closure of the efficiency gap by the end of the regulatory period) or a combination of two.

Direct translation of frontier efficiency scores into allowed revenue without adjustment or allowance of some time to reach 100% efficiency is not to be recommended. Any given X factor is a combination of backward-looking analysis of relative inefficiency and a projection of potential future frontier shift. Hence, it is inherently uncertain and open to legal

[42] See for example Agrell and Bogetoft (2018) who report on pan-European benchmarking exercises, which are not directly used in national X-factor setting. CRU (2025) reports undertaking frontier efficiency benchmarking as part of its recent price control review in Ireland, but not using in X-setting.

challenge if it is too onerous. Regulators who find they have been too generous in setting X for an individual firm can adjust this (using evidence of revealed outperformance) at the next price control period.

Benchmarking works best in circumstances where the firms are of similar sizes and have similar opportunities, and/or where firms are free to merge. If it were the case that small firms are at a substantial measured cost disadvantage and they were free to merge then the fact that a small firm was not capable of replicating the lower costs of a larger firm, might be less of an issue. However, merger costs are real and must be adequately accounted for.

Over time, any given benchmarking approach is likely to decline in efficacy. This is because benchmarking is likely to produce the strongest and most reliable guidance on relative efficiency (if implemented intelligently as above), in its earliest price control periods. However, over time any given method (or combination of methods) is likely to give rise to its own perverse marginal incentives given that it is simplifying the nature of efficient cost drivers.

Finally, energy regulators exist, inter alia, to benchmark the performance of regulated electricity distribution networks. However, regulatory use of benchmarking should evolve over time in line with best practice. Firm specific X factors should be more important in the earlier periods of the use of benchmarking, where there is both greater scope for efficiency improvement and greater variation in efficiency between firms. However, ‘simpler’ benchmarks (e.g. a common general productivity determined X factor with revenue sharing) can maintain incentives and may achieve the similar results for customers in the long-run. Certain regulators might therefore need to undertake a ‘benchmarking of benchmarking’ where their current increasingly complex and difficult to defend techniques are directly compared with simpler techniques to assess whether these might actually be equally or more effective.

References

AER (2024), *Annual Benchmarking Report Electricity distribution network service providers November 2024*, Commonwealth of Australia: Australian Energy Regulator.

Agrell, P.J., Bogetoft, P. and Mikkers, M. (2013). 'Smart-grid investments, regulation and organization', *Energy Policy*, 52: 656-666.

Agrell, P. and Bogetoft, P. (2018), 'Theory, Techniques, and Applications of Regulatory Benchmarking and Productivity Analysis', in Emili Grifell-Tatjé, C.A. Knox Lovell, and Robin C. Sickles (eds), *The Oxford Handbook of Productivity Analysis*, Oxford Handbooks, Oxford Academic, 10 Sept. 2018), https://doi.org/10.1093/oxfordhb/9780190226718.013.16

Agrell, P. and Bogetoft, P. (2014), *Review of Six Years of Energy Network Benchmarking*, Conference on benchmarking methods for incentive regulation in the energy sector: evaluating and improving the German approach, Bonn, 27/05/2014.

Agrell, P. and Bogetoft, P. (2007), PROJECT GERNER/AS6 Development of benchmarking models for German electricity and gas distribution FINAL REPORT 2007-01-01, Sumicsid AB.

Ajayi, V., Anaya, K.L. and Pollitt, M.G. (2022), 'Incentive regulation, productivity growth and environmental effects: the case of electricity networks in Great Britain', *Energy Economics*, 115 (November): 106354.

Ajayi, V., Anaya, K. and Pollitt, M. (2018), *Productivity growth in electricity and gas networks since 1990,* Energy Policy Research Group, University of Cambridge, Report prepared for the Office of Gas and Electricity Markets (OFGEM). Revised 21 December 2018. https://www.ofgem.gov.uk/sites/default/files/docs/2019/01/ofgem_productivity_report_dec_2018_1.pdf

Andor, M. and Hesse, F. (2014), 'The StoNED age: the departure into a new era of efficiency analysis? a Monte Carlo comparison of StoNED and the "oldies"(SFA and DEA)', *Journal of Productivity Analysis*, 41(1): 85–109.

Bauer, P.W., Berger, A.N., Ferrier, G.D. and Humphrey, D.B. (1998), 'Consistency Conditions for Regulatory Analysis of Financial Institutions: A Comparison of Frontier Efficiency Methods', *Journal of Economics and Business*, 50(2): 85-114.

Baumol, W.J. (1982), 'Productivity Adjustment Clauses and Rate Adjustment for Inflation', *Public Utilities Fortnightly*, July 22: 11-18.

Beesley, M. E., & Littlechild, S. C. (1989), 'The Regulation of Privatized Monopolies in the United Kingdom', *The RAND Journal of Economics*, *20*(3): 454–472. https://doi.org/10.2307/2555582

CC (Competition Commission) (2000), *Mid Kent Water plc: a report on the references under sections 12 and 14 of the Water Industry Act 1991*. London: Stationery Office.

CEER (2026), *Report on Regulatory Frameworks for European Energy Networks 2025, Incentive Regulation and Benchmarking Work Stream*, https://www.ceer.eu/wp-content/uploads/2026/01/RFR-2025-Main-report-combined.pdf

CEPA (2024a), *PR24 Draft Determinations: Frontier Shift, Real Price Effects and the energy crisis cost adjustment mechanism, Ofwat, 28 June 2024*. London: CEPA. https://www.ofwat.gov.uk/wp-content/uploads/2024/07/CEPA-frontier-shift-real-price-effects-and-the-energy-crisis-cost-adjustment-mechanism.pdf

CEPA (2024b), *EDB Productivity Study: A report prepared for the Commerce Commission 24 June 2024*, London: CEPA.

CMA (Competition and Markets Authority), (2022), Energy Licence Modification Appeals: Competition and Markets Authority Guide, CMA71, London: Competition and Markets Authority.
https://assets.publishing.service.gov.uk/media/635946898fa8f557cfc22c4f/Energy_Guidance.pdf

Coelli, T.J., Prasada Rao, D.S., O'Donnell, C.J. and Battese, G.E. (2005), *An Introduction to Efficiency and Productivity Analysis*, 2nd Edition, Boston, Kluwer.

Cronin, F.J. and Motluk, S.A. (2007), 'Flawed Competition Policies: Designing 'Markets' with Biased Costs and Efficiency Benchmarks', *Review of Industrial Organization,* 31: 43–67. https://doi.org/10.1007/s11151-007-9142-z

Cross, S., Padfield, D., Ant-Wuorinen, R., King, P. and Syri, S. (2017), 'Benchmarking island power systems: Results, challenges, and solutions for long term sustainability',

Renewable and Sustainable Energy Reviews, 80: 1269-1291.

CRU (2025), *Price Review Six: Distribution Revenue 2026 – 2030, Final Determination*, Commission for Regulation of Utilities.

Cunningham, M., Lawrence, D. and Fallon, J. (2020), *Frontier Shift for Dutch Gas and Electricity TSOs,* Report prepared for Netherlands Authority for Consumers and Markets

1 May 2020, Economic Insights Pty Ltd.

Europe Economics (2023), *Independent review of electricity benchmarking model for network companies*, DA4 Report 03 July 2023. https://ens.dk/sites/default/files/media/documents/2024-10/konsulentrapport_benchmarking.pdf

Eyre, S. and Pollitt, M.G. (2016), 'Introduction' in S.Eyre and M.G.Pollitt (eds.) *Competition and Regulation in Electricity Markets*, Cheltenham: Edward Elgar, pp.xiii-xxxv.

Frontier Economics (2023), *Benchmarking Techniques and Practices for Electricity and National Gas Network Operators, Report prepared for ACM 12 October 2023*, WWW.FRONTIER-ECONOMICS.COM

Greene, W. (2005), 'Reconsidering heterogeneity in panel data estimators of the stochastic frontier model', *Journal of Econometrics 126* (2), 269–303.

Haney, A. and Pollitt, M. (2013a), 'International Benchmarking of Electricity Transmission by Regulators: A contrast between theory and practice?', *Energy Policy*, Vol.62, November 2013, pp.267-281.

Haney, A. and Pollitt, M. (2013b), 'New Models of Public Ownership in Energy', *International Journal of Applied Economics*, Vol.27, No.2, pp.174-192.

Haney, A. and Pollitt, M.G. (2009), 'Efficiency analysis of energy networks: An international survey of regulators', *Energy Policy*, 37: 5814-5830.

Hesseling, D. and Sari, H. (2006), *The introduction of quality regulation of electricity distribution in The Netherlands*, European Energy Law Report III (EELR III-paer), 2006, pp.127-145. https://www.acm.nl/sites/default/files/old_publication/publicaties/11528_the-introduction-of-quality-regulation-of-electricity-distribution-in-the-netherlands-dennis-hesseling-mahir-sari.pdf

Jamasb, T. and Pollitt, M. (2008) "Reference models and incentive regulation of electricity distribution networks: an evaluation of Sweden's Network Performance Assessment Model (NPAM)." Energy Policy, 36(5): 1788-1801.

Jamasb, T. and Pollitt, M. (2007) "Incentive regulation of electricity distribution networks: lessons of experience from Britain." *Energy Policy*, 35(12): 6163-6187.

Jamasb, T. and Pollitt, M. (2001), 'Benchmarking and Regulation of Electric Utilities: Lessons from International Experience', *Utilities Policy,* Vol.9, No.3, pp.107-130.

Johnson, A.L. and Kuosmanen, T. (2015), 'An Introduction to CNLS and StoNED Methods for Efficiency Analysis: Economic Insights and Computational Aspects'. In: Ray, S., Kumbhakar, S., Dua, P. (eds), *Benchmarking for Performance Evaluation.* Springer, New Delhi. https://doi.org/10.1007/978-81-322-2253-8_3

Joskow, P.L. and Schmalensee, R. (2025), 'Cost-of-service regulation of electricity distribution services in the US' in Glachant, J-M., Joskow, P.L. and Pollitt, M.G. (Eds.), *Handbook on Electricity Regulation*, Edward Elgar, pp.44-67.

Joskow, P.L. (1974), 'Inflation and Environmental Concern: Change in the Process of Public Utility Price Regulation', *Journal of Law and Economics*, 17 (2): 291-327.

Kendrick, J.W. (1975), 'Efficiency Incentives and Cost Factors in Public Utility Automatic Revenue Adjustment Clauses', *Bell Journal of Economics*, 6(1): 299-313.

Kufeoglu, S., Pollitt, M. and Anaya, K. (2018), *Electric power distribution in the world: today and tomorrow.* Energy Policy Research Group Working Papers, No.1826. Cambridge: University of Cambridge.

Littlechild S. (2024), 'Incentive Regulation From the Inside: Resetting 12 RPI-X Price Caps in 1993–1995', *Review of Industrial Organization*, 65 (2): 375 - 430,

Littlechild, S.C. (2021), *An Overall Customer Satisfaction score for GB energy suppliers*, 12 March 2021, EPRG Working Paper 2027.

https://www.jbs.cam.ac.uk/wp-content/uploads/2023/12/eprg-wp2027.pdf

Littlechild, S. (2014), 'RPI-X, competition as a rivalrous discovery process, and customer engagement – Paper presented at the Conference The British Utility Regulation Model: Beyond Competition and Incentive Regulation?', *Utilities Policy*, 31: 152-161.

Littlechild, S.C. (1983), *Regulation of British Telecommunications' Profitability*, London: HMSO.

Llorca, M., Orea, L. and Pollitt, M. (2014), 'Using the latent class approach to cluster firms in benchmarking: An application to the US electricity transmission industry', *Operations Research Perspectives*, Vol.1 (1), pp.6-17.

Lovell, K. (2006). *Frontier Analysis: Recent Advances and Future Challenges*. Keynote Address, North American Productivity Workshop, Stern School of Business, New York University, New York.

Lucas, R.E. (1976), 'Econometric Policy Evaluation: A Critique' In Brunner, K.; Meltzer, A. (eds.). The Phillips Curve and Labor Markets. Carnegie-Rochester Conference Series on Public Policy. Vol. 1. New York: American Elsevier. pp. 19–46.
https://people.sabanciuniv.edu/atilgan/FE500_Fall2013/2Nov2013_CevdetAkcay/LucasCritique_1976.pdf

Meletiou, A. Vasiljevska, J., Prettico G., Vitiello, S. (2023), *Distribution System Operator Observatory 2022*, Publication Office of the European Union, Luxembourg, 2023 doi:10.2760/778963, JRC132379.

Montfoort van, F., Dijkstra, P.T. and Mulder, M. (2024), 'The impact of energy transition on distribution network costs and effectiveness of yardstick competition: an empirical analysis for the Netherlands', *Journal of Regulatory Economics*, 65: 85–107.

Nillesen, P. and Pollitt, M.G. (2026), 'Incentive Regulation and Distribution Network Performance: A Dutch Case Study (2000-2024)', *Journal of Regulatory Economics,* 69: 15. https://doi.org/10.1007/s11149-026-09511-5

Nillesen, P. and Pollitt, M. (2021), 'Ownership Unbundling of Electricity Distribution Networks', *Economics of Energy and Environmental Policy*, 10(1): 147-158. 10.5547/2160-5890.10.1.pnil

Nillesen, P.H. and Pollitt, M.G. (2011), Ownership unbundling in electricity distribution: empirical evidence from New Zealand, *Review of Industrial Organization*, Vol.38. No.1, pp.61-93.

Nillesen, P.H.L. and Pollitt, M.G. (2007) "The 2001-2003 electricity distribution price control review in the Netherlands: regulatory process and consumer welfare." *Journal of Regulatory Economics*, 31(3): 261-287.

O'Donnell, C.J. (2018), *Productivity and Efficiency Analysis: An Economic Approach to Measuring and Explaining Managerial Performance*, Springer.

Offer (1993), *Electricity Distribution: Price control, reliability and Customer Service*, October, Birmingham: Offer.

Ofgem (2024a), *Distribution System Operation (DSO) Incentive Report for Regulatory Year 1 April 2023 -31 March 2024*, London: Ofgem.

Ofgem (2022a), *RIIO-ED2 Final Determinations, Nov. 2022*, London: Ofgem.

Ofgem (2022b), *RIIO ED2 Draft Determinations Core Methodology Document, 19 June 2022,* London: Ofgem.

Ofgem (2022c), *RIIO ED2 Final Determinations Core Methodology Document, 30 November 2022,* London: Ofgem.

Ofgem (2010), *Regulating energy networks for the future: RPI-X@20 decision document*, London: Ofgem.

Olesen, O.B. and N.C.Petersen (1995), 'Chance Constrained Efficiency Evaluation', *Management Science*, Vol.41, No.3, pp.442-457.

Parameter, C.F. (2021), *Is it MOLS or COLS?,* Department of Economics, University of Oviedo, Efficiency Series Paper 04/2021. https://www.unioviedo.es/oeg/ESP/esp_2021_04.pdf

Pollitt, M.G., Covatariu, A. and Duma, D. (2024), *Towards a more Dynamic Regulation of Energy Networks*, Brussels: Centre on Regulation in Europe. https://cerre.eu/publications/towards-a-more-dynamic-regulation-for-energy-networks/

Pollitt, M. (2010), 'Does Electricity (and Heat) Network Regulation have anything to learn from Fixed Line Telecoms Regulation?', *Energy Policy*, Vol.38 (3), pp.1360-1371.

Pollitt, M.G. (2008) "The future of electricity (and gas) regulation in a low-carbon policy world." The Energy Journal, 29(S2): 63-94.

Pollitt, M. (2005) "The role of efficiency estimates in regulatory price reviews: Ofgem's approach to benchmarking electricity networks." Utilities Policy, 13(4): 279-288.

Pollitt, M.G. (1995), *Ownership and Performance in Electric Utilities.* Oxford: Oxford University Press / Oxford Institute for Energy Studies.

Rita, R., Marques, V., Bárbara, D., Chaves, I., Macedo, P., Moutinho, V. and Pereira, M. (2023), 'Crossing non-parametric and parametric techniques for measuring the efficiency: Evidence from 65 European electricity Distribution System Operators',

Energy, 283: 128511. https://doi.org/10.1016/j.energy.2023.128511.

Ros, A.J., Shetty, S. and Tardiff, T. (2024), 'Performance based regulation in electricity and cost benchmarking: theoretical underpinnings and application', Journal of Regulatory Economics, 66:62–94. https://doi.org/10.1007/s11149-024-09474-5

Shleifer, A. (1985), 'A Theory of Yardstick Competition', *The RAND Journal of Economics*, *16*(3): 319–327. https://doi.org/10.2307/2555560

Silva, H.C.D. (2007), *Empirical Essays in the Economics of Regulation*, PhD Dissertation University of Florida.

Trinkner, U. (2024), *Benchmarking of German Electricity DSOs: Experiences and Lessons Learned from RP2 to RP4.* BNetzA scientific dialogue, 24th September 2024, Presentation. Swiss Economics.
https://www.bundesnetzagentur.de/EN/RulingChambers/GBK/Calendar_old/09_2024/20240924_S1-1_Trinkner-Lessons_learned.pdf?__blob=publicationFile&v=2

Uri, N.D. (2002), 'A Note on Measuring the Productivity Offset for Regulatory Purposes', *Journal of Productivity Analysis* 17: 201–211.
https://doi.org/10.1023/A:1015055904013

Weisman, D.L. (2019), 'Regulatory Takings Without Confiscatory Returns', *Yale Journal of Regulation*, 36: 170-195.

Weyman-Jones, T. (2006), *Efficiency and productivity analysis in regulation and governance. 4th* NAPW Plenary Paper, New York.

Weyman-Jones, T.J. (2001), *Stochastic Non-parametric Efficiency Measurement in Electricity Distribution*, mimeo.

Yu, W., Jamasb, T. and Pollitt, M. (2009), 'Willingness-to-Pay for Quality of Service: An Application to Efficiency Analysis of the UK Electricity Distribution Utilities' The Energy Journal, Vol.30 (4), pp.1-48.

Appendix

Figure A.1 Example of regulatory of revenue profile over a five year price control from 2010-15

Appendix

Figure A.1 Example of regulatory of revenue profile over a five year price control from 2010-15