

PERFORMANCE BENCHMARKING IN UTILITY REGULATION: PRINCIPLES AND THE UK EXPERIENCE

David Parker
Cranfield School of Management
david.parker@cranfield.ac.uk

This Presentation is Based on the following forthcoming publication:

Parker, D., Dassler, T., and Saal, D. (forthcoming), "Performance Benchmarking in Utility Regulation: Principles and the UK Experience" in Cook, Paul, Crew, Michael, and Parker, David eds., Handbook on Economic Regulation. Edward Elgar.

BACKGROUND

- In the UK, the privatisation of utility industries has led to the development of regulatory regimes to prevent monopoly abuse
- Issue is: how to set price cap (or rate of return) while preserving efficiency incentives (productive efficiency)
- Littlechild (1983) provides virtually no guidance as to how price caps should be reset after the completion of a regulatory period
- One approach: base future revenue on the costs achieved by the firm in previous years, adjusted for any known exceptional items or expected cost changes.
- Incentives for cost efficiency?

BENCHMARKING

- Benchmarking (or ‘yardstick competition’): compare the performance of a regulated firm with some comparator.
- Shleifer (1985) argued that a regulated firm’s revenue needs should be assessed by looking at costs in comparable firms or industries. The objective should be to find ‘some relatively simple benchmark, other than the firm’s present or past performance, against which to evaluate the firm’s potential’.
- Shleifer’s solution was to identify ‘comparable firms to infer a firm’s attainable cost level’.
- Provided that the regulator has two or more firms under its jurisdiction, then yardstick competition can overcome information asymmetry.
- Each regulated firm, i , is assigned a ‘shadow firm’. The shadow firm becomes the benchmark for setting the revenue requirement.

- Benchmarking can provide regulators with information about efficient OPEX and CAPEX requirements, thereby reducing the regulated firm's informational rents.
- Benchmarking is now used extensively: for example in Costa Rica for transport tariff setting, in telecommunications regulation in Hungary, in Dutch electricity and telecommunicationS, and for electricity regulation in Norway and New South Wales
- Benchmarking costs reduces the effects of the company's own costs on prices; *in extremis*, if a firm's costs have no effect on its own revenues, the incentives for management to reduce costs will be maximised.

- 'It is essential for the regulator to commit to not paying attention to firms' complaints and to be prepared to let the firms go bankrupt if they choose inefficient cost levels. Unless the regulator can credibly threaten to make inefficient firms lose money... cost reduction cannot be enforced' (Shleifer, p cit., p.323).
- Shleifer was also aware that this form of yardstick competition required a 'shadow firm' comparable in terms of its cost structure. The result would also be unreliable if the firms faced different demand functions; although for simplicity Shleifer assumes a common demand function for much of his analysis.
- To overcome the lack of a perfect 'shadow firm', Shleifer recognised that multivariate regression models would need to be developed to reflect characteristics that could account for cost differences between firms that are not within the control of management (e.g. topography, customer density, regional wage costs, etc.).

- Shleifer warns against the marginal cost of the exercise overwhelming its marginal benefit in terms of more accurate price setting.
- He makes reference to the dangers of ‘collusive manipulation’ of yardstick competition by participating firms. Yardstick competition requires managerial independence across the firms used as comparators; otherwise the results will be biased

METHODS USED

- ***Productivity Indices,***
- ***Stochastic analysis of production and cost functions, such as translog cost functions,***
- ***Mathematical modelling, especially the use of DEA (data envelopment analysis).***
- ***Engineering models.*** An alternative approach is to calculate theoretical production functions based on engineering data.
- **Bauer et al. (1998)** have proposed a set of consistency conditions when different measures are adopted by regulators, broadly the different methods used should provide consistent efficiency levels and rankings and identification of best and worst performers, and also be consistent in their results over time.
- **Coelli and Perelman (1999)** have suggested combining the results from alternative modelling exercises by using the geometric means of the performance scores for each data point in order to reduce potential bias. [

The UK Experience

- In the UK, benchmarking performance is now part of resetting regulatory price caps and is used to help determine appropriate X factors.
- Over time there appears to have been movement towards a more common approach to benchmarking across the regulatory offices - the result of increasing experience, Monopolies and Mergers Commission (now Competition Commission) investigations, demonstration effects and government policy that favours more consistency in regulation across the different regulatory bodies (DTI, 1998, para.80). Nevertheless, some important differences remain.
- The regulatory offices undertake their own productivity and cost analyses and from time to time use research produced by (different) outside consultancy firms.
- Expert judgement also plays a part in efficiency assessment by the regulatory offices and the experts come from diverse backgrounds including industry and academia.
- There is, therefore, some heterogeneity in the approaches adopted by UK regulatory offices when setting price caps.

TELECOMMUNICATIONS

- **At privatisation BT's price cap was based on discussion between the government, the company and the City.**
- **When OFTEL came to reset the price cap from the late 1980s, it began to experiment with benchmarking. Service baskets were constructed for the UK, France, Italy and West Germany, which are the four largest telecoms markets in Europe. The baskets included the number of calls by distance, time of day, day of week and duration. However, OFTEL acknowledged the difficulties involved in making useful comparisons, especially because pricing differed substantially between countries. OFTEL continued to experiment with international benchmarking in the following years, because within the UK there was no available comparator to BT.**
- **Despite the existence of international benchmarking, it is unclear how it influenced the setting of BT's X factor and its service targets, if at all, in the early years.**
- **The main influence on price cap setting seems to have been BT's own past and forecast cost performance. This, of course, rules out the establishment of an efficient frontier and the application of catch-up efficiency targets determined by such a frontier**

- **The use of benchmarking was more evident after 1995, when an international benchmarking exercise was commissioned from NERA. In this study BT's OPEX and CAPEX levels were separately compared to those of Local Exchange Carriers in the US. BT's interconnection charges were also benchmarked against those of domestic mobile networks and international ISDN charges. The NERA study was later complemented by an international comparison of corporate telecommunication costs and by an international comparison of BT's interconnection charges.**
- **In the run-up to the most recent price control review for BT, in June 2003, OFTEL again used international comparisons, for the Internet, fixed line and mobile markets. The range of prices within these markets in a number of countries was compared, alongside indicators of service quality.**
- **NERA applied OLS, SFA and DEA methods to derive comparative measures. The results from these exercises were complemented by OFTEL's own performance estimates, including figures for market shares and domestic market growth within the UK. An interesting result of the exercise was the conclusion that UK interconnection charges and call tariffs were uncompetitive.**

WATER & SEWERAGE

- From the outset OFWAT placed considerable emphasis on yardstick competition in price cap setting. Indeed, its use is legislated for in section 34(3) of the Water Act 1991.
- given the relatively large number of companies in the water sector benchmarking performance through econometric means was perceived to be a viable option at privatisation.

- **In the mid to late 1990s OFWAT seems to have abandoned the use of DEA entirely and relied on econometric modelling.**
- **Water supply CAPEX was treated separately and again water and sewerage were considered independently. The 1999 price review used four models for water OPEX and these were similar to (though not the same as) those used in 1994.**
- **The results from the different econometric models were summed to create efficiency bands, after making adjustments for company-specific costs considered to be outside the control of management (e.g. regional salaries, requirements for additional water treatment and where there are large numbers of small water sources).**
- **An allowance was also made where companies were assessed as efficient in OPEX and inefficient in CAPEX, or vice versa. This was undertaken to address criticism from the industry that modelling OPEX and CAPEX separately risked inappropriate efficiency scores.**
- **New in 1999 was the application of a price cap adjustment to reward or penalise service quality.**
- **OFWAT then set the price caps for 2000 to 2005 on the basis that 60% of the OPEX gap would be closed over the five years (OFWAT, 1998a, 1998b, 1999a, 1999b).**

- Using the efficiency bands, OFWAT determined CAPEX in terms of catch-up and frontier efficiency movements for capital maintenance and capital enhancement schemes.
- For both the companies' own cost assumptions formed a starting point, again based on their unit costs for a range of specimen projects. These figures were then subjected to expert assessment by consulting engineers. They were then compared and adjusted for each company to reflect OFWAT's judgment of the efficiency of each company's costs relative to the industry as a whole.
- For capital maintenance, four econometric models covering resources and treatment, distribution infrastructure, distribution non-infrastructure, and management and general expenses were also used. The data in the models were expenditures over the five years from 1993/4 to 1997/8, to even out year to year spending fluctuations

- Over time, based on its modelling exercises, OFWAT appears to have determined the efficient frontier in different ways.
- In 1994 the benchmark for capital schemes was the lower quartile of the cost population for each standard cost.
- In 1998 this was changed to the lowest reported standard costs for groups of standard costs meeting certain criteria.
- In 1994 catch-up was calculated based on 50% of the difference between each company's reported standard cost and the benchmark cost.
- In 1999 this was set at 50% for capital maintenance and 75% for capital enhancement, and now the catch-up was to be achieved in the first year of the price cap period and not over its full five years as previously.

DISCUSSION: ISSUES

- The accuracy of the resulting rankings
- The limited use of international comparisons
- Treatment of the quality of service
- The regulatory burden
- Uncertainty

CONCLUSIONS

- **In 1985 Shleifer proposed the use of benchmarking or yardstick competition in economic regulation so that a regulated firm's prices were not based on its own costs.**
- **In the UK benchmarking is now part of the process of setting price caps. OFTEL, OFGEM and OFWAT have experimented with both national and international benchmarking, although it appears that less importance has been placed on international data because of the difficulty of finding acceptable comparators. The main exception is in telecommunications where BT's continued dominance in the market restricts the use of domestic comparator firms.**
- **The use of imperfect benchmarking may be better than relying only on cost information supplied by the regulated firm. At the same time, however, the operation of benchmarking has proved to be far from problem free in the UK.**

- **It has been subject to criticism relating to (1) the selection of the appropriate performance model; (2) the accuracy of the resulting efficiency scores; (3) the extent with which international benchmarking is used; (4) building in correct allowances for service quality; (5) the costs imposed on the industry ; and (6) the uncertainty that is said still to surround the price capping exercise. As Burns et al. (2005) comment, to be successful benchmarking needs to be seen as having high value and credibility, and in the UK both appear to be lacking, with the partial exception of in water industry regulation.**

- Nevertheless, if we accept both that that the natural monopoly network components of infrastructure industries will continue to require regulation, and that the benefits of price cap regulation remain sufficiently strong to rule out the use of rate of return regulation, benchmarking must be seen as an imperfect but necessary tool in infrastructure industry regulation.
- Thus, while the UK experience of benchmarking has proved more complex and subject to error than originally anticipated, it also demonstrates that it has been an important input into the regulatory process.