
The Appropriate Role of Yardstick Methods in Regulation

Brian Williamson
Susanne Toft

May 2001

n/e/r/a

National Economic Research Associates
Economic Consultants
<http://www.nera.com>

CONTENTS

Executive Summary.....	1
Yardstick Competition in Context.....	3
Origins of Yardstick Competition in the Water Industry.....	4
Incentive Properties of Yardstick Competition	5
Practical Problems in Applying Yardstick Competition.....	8
The Appropriate Role of Yardstick Methods.....	13
Appendix A: Theoretical Underpinnings of Yardstick Competition.....	16

Executive Summary

We were commissioned by Northumbrian Water Limited to consider the appropriate role of yardstick methods, and yardstick competition in particular, in the regulation of the water industry in the UK. We use the term yardstick competition to refer to the use of comparisons (statistical or otherwise) by a regulator to set a company's allowed revenues.

Yardstick Competition in Context

Competition, where feasible, represents the ideal “yardstick method”. The market price provides a direct yardstick since companies compete *in the same market*, and companies that provide a product or service consistent with customer preferences at least cost will prosper. In contrast, for regional monopoly water companies, costs or quality may differ for perfectly legitimate reasons (since companies operate in *different markets* with different environments and customers). There is no direct measure of relative performance *in the same market* – and regulation cannot hope to mimic competition.

In a monopoly market one possible benchmark is an individual company's actual costs – although the realisation that cost reduction will lead to price reduction will blunt incentives for efficiency. The aim of price caps is to partially restore incentives and strike an appropriate trade-off between incentives for efficiency and the transfer of benefits to customers over time.

The aim of yardstick competition is to improve on outcomes achievable under price caps by sidestepping the trade-off by setting prices in line with an “efficient” cost standard determined by other companies costs. However there are practical and theoretical constraints on what yardstick competition can be expected to achieve that raise a question over whether outcomes under price caps can be improved on, or whether attempts to apply yardstick competition will in practice harm incentives for efficiency.

Problems in Applying Yardstick Competition

Shleifer set out the necessary conditions for yardstick competition to promote efficiency in 1985 (discussed in Appendix A). He noted that where firms operate in identical environments “*It is essential for the regulator to commit himself not to pay attention to the firms' complaints and to be prepared to let the firms go bankrupt...*”; and that where operating environments differ (and such differences are not fully accounted for) a subsidy may be required “*to run the scheme without bankruptcies*”. Either alternative is problematic – raising a question over the benefits of yardstick competition in practice.

Yardstick competition may lead to efficient firms being bankrupted by “mistake”, or to assistance to inefficient firms “*to run the scheme without bankruptcies*”. Once one envisages bringing prices more closely into line with actual costs (and away from yardstick costs) the claimed advantages of yardstick competition over price caps or cost plus regulation are eroded. Indeed, yardstick competition may actually undermine efficiency.

The application of yardstick competition is inherently subjective in practice since there is no objective basis for attributing unexplained cost differences

to *inefficiency per se*, or deciding over what period company specific “catch-up” should occur. In turn, regulatory discretion (and potential for opportunism) undermine incentives by reducing confidence that efficiency savings will be retained, and by undermining assurance of cost recovery for sunk investments.

Regulatory discretion may also raise the cost of capital. Comments by Moody’s in February 2001 suggest that yardstick competition is perceived as a key source of regulatory uncertainty:

“The main areas of uncertainty associated with this method of regulation [incentive based regulation] arise from the principle of comparative competition applied by Ofwat in setting opex and capex allowances...”

Preserving individual companies as yardstick comparators may prevent efficiency gains from structural adjustment, and reduce capital market pressures for efficiency on management.

The Appropriate Role of Yardstick Methods in Regulation

While we argue that yardstick competition is flawed as a regulatory method, comparisons can provide useful information and therefore have a role to play in both management and regulation. Management within companies, owners, and market analysts can and do apply benchmarking techniques to identify and/or reward efficiency. For example, senior managements’ pay may depend in part on other companies’ performance. Provided the regulatory regime offers incentives for efficient behaviour companies themselves can be expected to apply benchmarking techniques where appropriate.

Under price cap regulation a basis for setting the “X” factor in “RPI-X” formulae is required. One approach is to use past industry productivity *trends* as a basis for setting expectations. This represents a form of benchmarking – though it differs fundamentally from yardstick competition in that there is no need to normalise cost *levels*, and the starting point for the application of the “X” factor is actual company specific costs.

Regulatory parameters that are not directly observable may need to be estimated using industry wide data, for example, the cost of equity finance. However, in estimating such parameters judgement should be applied within narrow and well-defined bounds using agreed methods (for example, the capital asset pricing model).

Various kinds of benchmark may be used to identify expenditures or areas of service quality performance that justify further scrutiny. However, unexplained cost differences are not assumed to represent *prima facie* evidence of *inefficiency*, and price caps operating in the more stable and predictable environment proposed would provide stronger incentives for utilities to *reveal* efficient cost levels over time (rather than regulatory reliance on benchmarks to decide what costs *ought to be*).

Yardstick Competition in Context

We use the term *yardstick competition* to refer to the regulatory use of comparisons (statistical or otherwise) to set a company's allowed revenues on the basis of other companies' costs. *Yardstick methods* in regulation, however, involve a wide range of techniques and applications aside from yardstick competition. Yardstick methods may be applied in assessing cost or service quality levels, or in assessing future expectations (forecasts) of cost or service quality. Finally, a range of parties including regulators, company management, and capital market participants may also apply yardstick methods.

We focus primarily on the question of the appropriate role of yardstick methods in setting allowed revenues. In assessing alternative yardstick methods we assess yardstick competition relative to incentive based regulation whereby incentives are offered to companies to *reveal* efficient cost levels over time, rather than the regulator attempting to assess what efficient costs *ought to be*.

Actual competition is the ideal yardstick "method" - where feasible. Competition ensures that incentives for efficiency are strong since firms in a competitive market face a strong test of their success or failure – *their performance relative to others in the same market*. No statistical normalisation or other adjustment is required to allow for underlying differences in operating environment in assessing performance since firms are free to choose their location (within planning constraints), and operate as price takers facing a common market price.

Competition can achieve both productive efficiency (via strong incentives for cost reduction), and allocative efficiency (via a market price which constantly adjusts in line with costs). However, where competition is not feasible or efficient for network utilities a trade-off arises between setting prices in line with costs (preventing monopoly abuse) and offering incentives (profit opportunities) for companies to improve their efficiency over time. Price caps, by introducing a fixed lag between cost and price reduction, aim to strike an appropriate balance between incentives (long term customer interests) and immediate price reduction. Yardstick competition may be viewed as an attempt to mimic competition, by sidestepping the regulatory trade-off involved in basing prices on actual company specific costs.

In a spectrum of benchmark methods "yardstick competition" is ambitious in its objective, and demanding in terms of data and analysis (a range of other benchmarks or yardstick methods might also be applied by management, capital markets, and regulators that are less ambitious and demanding). *A like with like* comparison of performance *in the same market* is not possible for regional monopolies since costs and quality will differ for legitimate reasons due to variations in operating environment and customer base. In practice, therefore, yardstick competition cannot mimic competition. Regulators do, and arguably should, consider actual costs in setting allowed revenues.

In an effort to normalise costs and allow a like with like comparison the regulator may use simple ratios or statistical (econometric) methods. However, normalisation will inevitably be imperfect, and those

companies disadvantaged relative to actual costs (or facing insolvency) may challenge the regulator's decision. The additional demands of yardstick competition relative to other forms of benchmarking were recognised early on in relation to the water industry:¹

“there is a considerable difference between the degree of comparability needed for a qualitative comparison of performance, and that needed for an objective basis for regulation which will stand up to scrutiny and (ultimately) legal challenge.”

Disputes may therefore be more likely under yardstick competition since allowed revenues may diverge significantly from actual costs under this methodology, cost normalisation is a very inexact science, and the consequences of “mistakes” are potentially large. In addition, the transparency of yardstick competition in the water industry has been low – in particular it is not possible to replicate the approach adopted by Ofwat including the translation of normalised costs into allowed revenues on a company-by-company basis.

Independent of the role of yardstick methods in regulation private companies, their owners, and capital market participants may use benchmarks to assess the performance of companies. Private or “voluntary” benchmarking tends to focus on process issues, and identifying exceptional costs or performance – rather than the use of statistical methods to set “budgets” for activities. Benchmarks may be used in assessing performance pay for individuals.

Short of yardstick competition regulators may nevertheless find it necessary or desirable to rely in part on benchmarks of various kinds. For example, regulatory parameters such as the market cost of equity or future costs do not appear in the accounts and must be estimated in some way, perhaps using industry benchmarks.² However, such applications of benchmarking should be carried out in accordance with an agreed method (for example, the capital asset pricing model), and do not involve the wide range of discretion and judgement that is arguably inherent to yardstick competition involving total cost benchmarking.

It is not necessary for the regulator to know the “efficient” level of costs – indeed the premise of incentive based regulation is that the efficient level of costs is not known and can only be revealed over time in response to appropriate incentives. However, baseline assumptions are nevertheless required around which incentives can operate. In price cap plans the purpose of the “X” factor is therefore to pass *anticipated* savings to customers. A benchmark based on past rates of productivity growth is one way to establish a value for “X”.

Origins of Yardstick Competition in the Water Industry

In his 1986 report on the regulation of privatised water authorities Littlechild, drawing explicitly on the article by Shleifer (1985) considered in Appendix A, proposed an industry wide yardstick:³

“The different conditions of the ten authorities initially suggest the need for up to ten different values of X. However, there is a stronger case for a uniform X across all authorities. This would avoid accusations of discrimination, facilitate operation of the ‘industry yardstick’ for revising X, and minimise

the burden of licence negotiations.”

Littlechild therefore proposed an approach consistent with using industry wide rates of productivity growth to set “X” (cost level benchmarking does not yield an estimate of “X”). The practice of setting “X” based on industry wide Total Factor Productivity (TFP) trends has emerged as best practice in the US.⁴ However, subsequent discussion in the UK of yardstick competition focussed on benchmarking cost levels.

Over the period December 1993 to February 1994 Ofwat published three research papers written by Professor Mark Stewart on the impact of operating conditions on (water and/or sewerage) company costs using econometric modelling, and at the 1994 Review carried out ordinary least squares (OLS) analysis of company’s overall operating costs. Data Envelopment Analysis (DEA) was used in addition as a robustness check.

At the 1999 Price Review OLS analysis was carried out on operating costs (opex) and capital maintenance costs. The assessment of relative efficiency for operating costs was mainly based on econometric models, while the assessment for capital maintenance was based on econometric models and unit cost comparisons. Modelled operating costs excluded various “uncontrollable” costs.

After taking account of company specific factors the results from the econometric models and unit cost comparisons were used to rank companies. Companies were then divided into performance bands that were used to set the “catch-up” factors for each company. The approach adopted differed from that originally suggested by Shleifer in a number of important respects including the fact that a “frontier” company was used to assess catch-up factors (rather than a simple average or mean OLS regression) and the results were used to set “X” factors in a price cap formula. The price cap formula operates for five years between reviews, at which point “X” and other factors are reassessed. To avoid periodicity in relation to incentives for cost reduction a rolling five year outperformance mechanism, whereby savings are retained for five years irrespective of their timing, was implemented by Ofwat in 1999.

Incentive Properties of Yardstick Competition

In this section we consider the incentive properties of yardstick competition, and the impact of yardstick competition on the cost of capital. It is argued that incentives (and constraints on management) are the underlying determinants of efficiency over time, and that yardstick competition may harm rather than improve incentives for the following reasons:

- i. The assumptions underpinning the theoretical efficiency properties of yardstick competition are unlikely to be met in practice (either because they are too demanding of empirical modelling, and/or because they are inconsistent with rational conduct).
- ii. In practice a purely exogenous benchmark is not applied – particularly for company chosen to represent the hypothetical “efficiency frontier”.
- iii. The methods of yardstick competition are inherently subjective (and

discretionary) and regulatory discretion (and potential opportunism) undermines incentives and raises the cost of capital.

- iv. Efforts to preserve yardstick comparators inhibit potential efficiency gains from structural adjustment and reduce pressure on management.

Focus on Efficiency or Incentives?

In practice, applications of yardstick competition tend to focus on identifying whether companies are efficient or not – rather than on improving incentives. As Kahn put it in relation to hypothetical models of incremental costs (so called TSLRIC) in US telecommunications “*the Commission [FCC] has in effect declared, ‘We will determine not what your costs are or will be but what we think they ought to be.’*”⁵ (The Federal Appeals Court ruled against the use of TSLRIC in July 2000.)

Identifying “inefficiency” is not a sensible goal of regulation judged against widely accepted principles of economics that imply that “inefficiency” (relative to an unattainable world of perfect information) is likely to be prevalent in market economies. The relevant question for regulation is therefore not to decide what an efficient outcome would be, but whether the incentives on offer can be improved (or whether unnecessary constraints on owners or management can be removed).

As Weisman noted:⁶

“the appeal of the efficient-firm cost standards is fleeting once it is recognized that the logic on which it is based is inherently circular. If regulators were in fact all-knowing, there would be no need for competition. The regulator could simply direct the incumbent firm to produce in accordance with the efficient-firm cost standard. The reality, of course, is that regulators do not have sufficient information to actively engage this approach. The pervasive adoption of price cap regulation is confirmation of this very fact.”

The Theoretical Efficiency Properties of Yardstick Competition Rest on Unrealistic Assumptions

The theoretical incentive and efficiency properties of yardstick competition are grounded on assumptions that were set out by Shleifer (1985) raising the possibility of applying yardstick methods to regulation. However, these requirements are either unlikely to be met in practice, inconsistent with rational conduct, or would reintroduce the incentive problems yardstick competition is intended to overcome (see Appendix A).

Essentially, the theoretical efficiency properties of yardstick competition rest on a willingness to bankrupt firms operating in identical environments (or firms where cost differences due to differences in environment have been fully accounted for). However, committing to bankrupt firms judged “inefficient” might be neither credible nor rational (given the likely costs involved and/or legal constraints on the conduct of the regulator). Shleifer discussed the possibility of subsidising firms that would otherwise be bankrupted. However, subsidies would reintroduce the incentive problems yardstick competition is intended to

overcome ie endogeneity of allowed revenues and actual costs.

“Yardstick” Revenues Have Been Endogenous in Practice

In practice, in the UK water industry, yardstick competition is combined with price caps in a hybrid system whereby yardstick evaluations enter a judgement about what “X” in the “RPI-X” formula should be for each company. Companies allowed revenues are then at least in part dependent on their actual costs – in particular for a company that expects to be on the efficiency “frontier” over time.⁷

Yardstick Competition Is Inherently Subjective & Uncertain

Yardstick competition is inherently subjective since judgements are required on what explanatory factors to include, the model to adopt, and the translation of model outcomes into assumed efficiency differences and rates of “catch up”. The exercise of subjective judgements under yardstick competition in turn introduces uncertainty for companies, and introduces scope for regulatory opportunism and retrospective conduct that undermine incentives. The practice of yardstick competition in the water industry lacks transparency - thereby increasing uncertainty and reducing regulatory accountability.

Non-transparency, discretion, and potential opportunism undermine incentives for a number of reasons:⁸

- Companies do not know to what extent they will benefit from cost savings.
- Irreversible investment in long-lived assets is discouraged by regulatory discretion since companies do not have a reasonable assurance of cost recovery.
- Companies may respond by “biasing” operating and capital investment decisions based on assumptions regarding the regulators conduct.

Uncertainty *per se*, as opposed to opportunism and retrospection that can flow from the exercise of undue discretion, may raise investment hurdle rates due to real options effects; make it harder for owners or senior managers to monitor and reward good management; and increase the risk that free cash flow is inadequate to fund socially worthwhile investments.

Uncertainty introduced by yardstick competition may also raise the cost of capital. Evidence that yardstick competition is a key source of regulatory uncertainty is provided by comments by the rating agency Moody’s on the risks in the water industry:⁹

“Under this regulatory regime [incentive based regulation], water companies enjoy the protection of what is essentially a mechanism for the re-alignment of both business and financial costs with prevalent market prices at a maximum of five-yearly intervals. The main areas of uncertainty associated with this method of regulation arise from the principle of comparative competition applied by Ofwat in setting opex and capex allowances, and from the discretion

that the Regulator retains in assessing the industry's unit costs and cost of capital within the statutory framework and the stated methodology."

Preserving Comparators May Harm Efficiency

Finally, we note that efforts to preserve comparators for yardstick competition may reduce the scope for efficient structural change and reduce pressure on management to be efficient. Littlechild, commenting on the proposed approach to the regulation of the water industry in 1986, emphasised the importance of capital market pressure in promoting efficiency:¹⁰

"[Concern over]...management slack and empire-building can be met by fostering competition in the capital market – in particular, by maintaining the threat of takeover."

Professor Allan Fels, Chairman of the Australian Consumer Competition Commission, expressed scepticism regarding merger restrictions based on a desire to preserve a threshold number of companies:¹¹

"I would agree with those that argue that while yardstick competition is of assistance to regulators in assessing the performance of natural monopoly suppliers, it would not fully replicate market pressures. In other words it does not equate to a market. Therefore, to give it market attributes, such as minimum threshold numbers of participants, and make maintenance of such thresholds a pivotal reason for rejecting mergers, seems somewhat off target"

The analysis in this section suggests that comparative competition involves efficiency costs that may outweigh the benefits, even before account of potential foregone merger savings. In addition, structural changes in the water industry, including greater use of competitive procurement, raise further doubt over the current emphasis on the preservation of separate companies for the purpose of comparative competition.

Practical Problems in Applying Yardstick Competition

Problems in relation to the application of yardstick competition in the water industry have been noted in the past.¹² In practice, problems in applying yardstick competition in the water industry arise at each of the following steps:

- i. Allowing for exogenous cost drivers using modelling to determine residual unexplained cost differences;
- ii. Judging whether unexplained cost differences are due to random errors or efficiency differences; and
- iii. Translating assumed differences in "efficiency" into productivity growth assumptions ("X" factors).

The inherent limitations of data and modelling therefore enter the estimation of unexplained cost differences, while judgement necessarily enters the assumptions regarding company specific "efficiency" differences and their translation into productivity growth assumptions. In practice the steps are not set out in published documents with

sufficient clarity to allow the results for each and every company to be reproduced independently.

While some company specific information is made available to the individual company concerned, an overall appraisal of the method, and meaningful ability to challenge the assumptions, is limited by a lack of information regarding the treatment of all companies. In addition, the approach to future reviews is unclear. Companies are therefore unsure as to how company specific catch up factors have been, or will be, determined (and therefore of the net gains from outperforming past efficiency targets), and whether recovery of prudently incurred costs will be allowed in future.

Allowance for Exogenous Cost Drivers is Incomplete

Not all cost drivers in the water industry are precisely known or adequately defined, those that are known cannot necessarily be objectively measured, and models of “unexplained” cost differences do not include all known and measured cost drivers.¹³ The modelled relationship therefore provides an incomplete explanation of cost differences due to exogenous factors. For example, commenting on the Ofwat distribution costs model Professor James Davidson noted:¹⁴

“In this case, the single included variable is the ratio of mains exceeding 300mm diameter to the total mains length. The proposition that this model completely explains cost appears, on the face of it, incredible.”

The difference between actual and predicted costs in a regression model picks up the influences on costs of any relevant variables that have not been included in the regression as cost drivers. Therefore not only do modelled costs provide an incomplete explanation of cost differences, the estimates for each company may be biased.¹⁵

Output quality is one example of an important variable that is a cost driver, but is not included. Likewise differences in regional wages is an example of a variable recognised as important, but not included on grounds that an appropriate measure was not available. Prior to the 1994 review Professor Stuart noted that:¹⁶

“Ofwat’s view is that there is no available measure of regional wage variation that is both relevant to the water industry and derived on the basis of information from competitive industries”.

A number of other “company specific” adjustments were made after the econometric modelling was completed – though the magnitude of these was not published.¹⁷ Other evidence points to the fact that not all exogenous factors were captured. In particular, model results have not proved robust to plausible model and data set changes.¹⁸

Identifying Efficiency Differences (vs. Unexplained Cost Differences)

Utilities’ costs differ for reasons unrelated to efficiency differences, and it will be impossible to establish whether *unexplained* cost differences are due to differences in efficiency or other factors. Commenting on the

techniques of efficiency assessment adopted by Ofwat, Dr Driffield and Professor Davidson noted:¹⁸

“As practising econometricians, we feel that the rationale for treating the residuals as indicators of efficiency is doubtful at best. This is particularly the case where the dataset is incomplete, lacking important cost drivers, or where there are inconsistencies in the way that the data have been collated by different companies. Given the specification uncertainty, we would expect a regulator to treat a company as inefficient only if it shows up as a significant and robust outlier over alternative specifications, and if that judgement is confirmed by detailed investigation “on the ground”.

Professor Mark Stewart made the same point in his work for Ofwat prior to the 1994 Periodic Review:¹⁹

“It should be noted that the inefficiency of a company is an inherently ‘residual’ concept. We control for what cost drivers we can. After this some companies have higher costs (given the drivers) than others and we call this inefficiency. The accuracy of the measures derived therefore depends on the extent to which we are able to control for all relevant cost drivers.”

In practice the “decomposition” of unexplained cost differences into differences due to random errors and efficiency *per se* necessarily involves subjective judgement.²⁰

Translating “Efficiency” Differences into “X” Factors

Ofwat set out their approach to translating “efficiency” differences into “X” factors as follows:²¹

“The current preferred approach is to use the leading comparator as the benchmark, provided that this represents a reasonable proportion of the industry, and then phase the catch-up towards this performance over the price limit period.”

However, several problems arise in relation to the translation of estimated efficiency differences into company specific “X” factors. First, Ordinary Least Squares (OLS) regression provides an estimate of the mean of a distribution, not a “frontier”. Second, cost benchmarks for expenditure sub-categories, benchmarks at different points in time, and benchmarks involving different degrees of risk sharing with contractors are combined and may not be commensurate. Third, an assessed cost difference due to “inefficiency” does not directly yield an “X” factor.

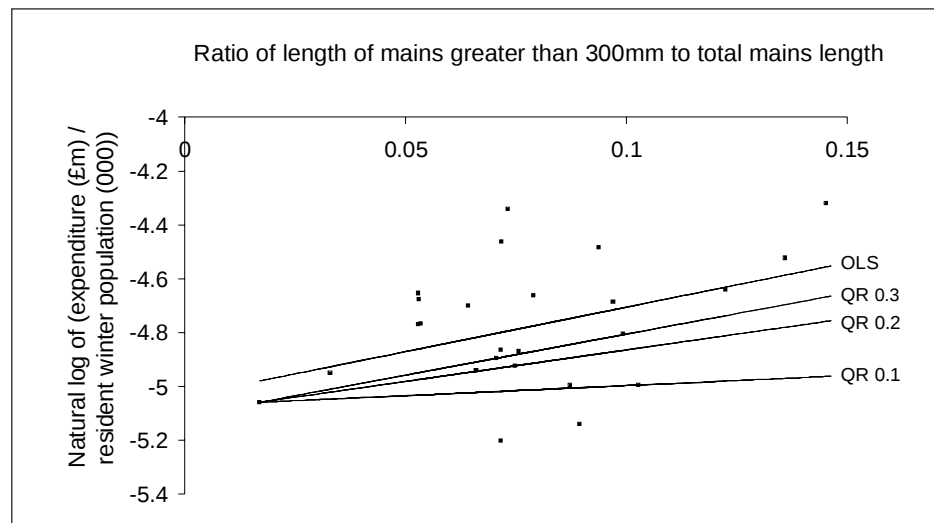
Regression generates an estimate of mean costs – not a “frontier”

Running an OLS regression provides an estimate of unexplained cost differences relative to the industry mean. In practice for the water industry the OLS line is then effectively shifted parallel to itself to provide an estimate of the cost “frontier”. Company specific “catch-up” factors are assessed relative to this frontier. However, adjusting the intercept, but leaving the initial OLS regression estimates of the slope parameters unchanged, does not necessarily provide an appropriate estimate of a (chosen) hypothetical “frontier” in the distribution.

An alternative method would be to use quantile regression to estimate unexplained cost differences directly at a given point in the distribution (for example, relative to the 20th percentile of costs). Quantile regression may produce slope coefficients that vary from those obtained using OLS regression.²³ Unexplained cost differences for individual companies will thus differ between a quantile regression and an adjusted OLS estimate.

Figure 1 illustrates how the slope of the OLS and various quantile regression lines can differ. Quantile regressions are shown for the 10th, 20th and 30th cost percentiles.

Figure 1
OLS & Quantile Regressions for Water Distribution Expenditure



Source: NERA analysis using Ofwat data for 1999/00.

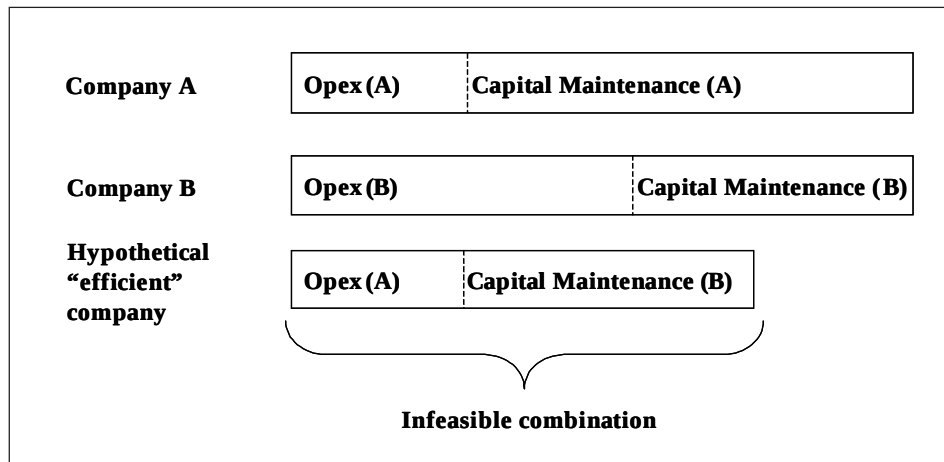
The robustness of “frontier” and median or OLS estimates will also differ:²⁴

“Furthermore, “best practices” are much more difficult to estimate accurately than, say, average or median performance. Indeed, the rationale for incentive-based regulation is grounded in the idea that the regulator cannot estimate minimum costs especially accurately.”

Benchmarked sub-categories may sum to an infeasible benchmark

A potential problem in combining modelled cost sub-categories (say operating costs and capital maintenance) is illustrated in Figure 2. Such differences may arise due to differences in inheritance (which could persist for very long periods of time in the water industry); due to differences in accounting treatment of costs (which can be minimised but never eliminated), and differences due to efficient combinations of inputs differing across companies (for example, due to regional differences in relative input prices for capital and labour).

Figure 2
Infeasible Combinations of Costs

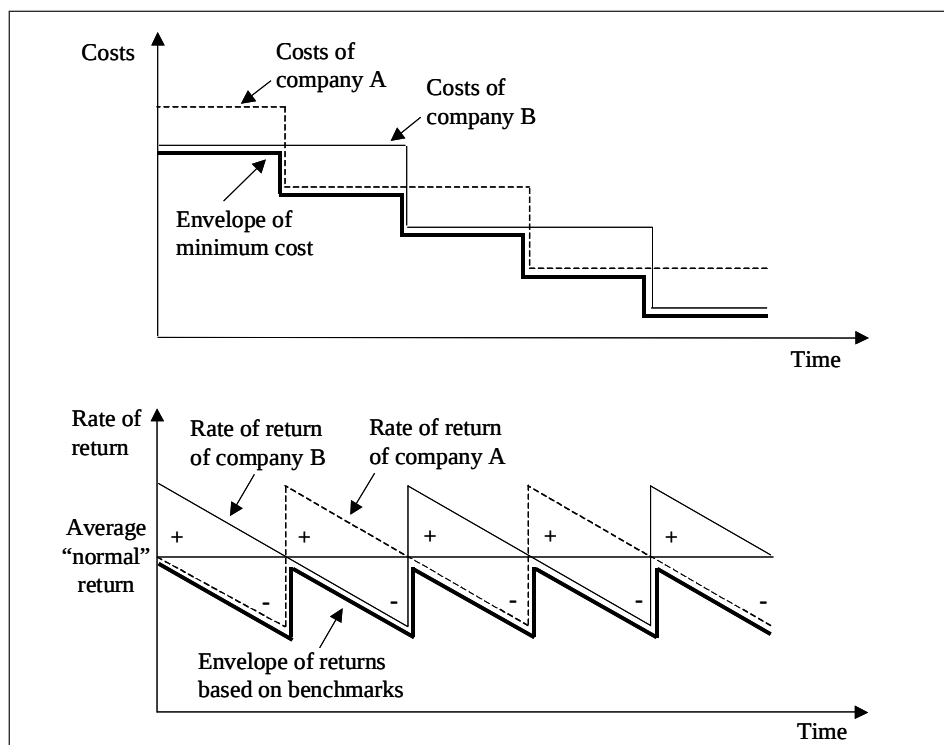


Unaccounted for differences will lead to a systematic downward bias of the combined cost estimate. Cost benchmarks may therefore reflect an infeasible “hybrid” or “model” company rather than achievable performance for any real company.

An envelope of benchmarked costs may prove infeasible over time

Adjustments of inputs such as labour or capital tend to be lumpy with discrete changes at different points in time.²⁵ Returns will then fluctuate, and each company will have periods of above and below “normal” returns. Figure 3 illustrates this situation – and the possible impact of cost benchmarking.

Figure 3
Benchmark Based on Minimum of Companies' Costs Over Time



If each company is constrained to the minimum cost envelope at each point in time, then the cost envelope and/or actual returns will be inadequate.

Differing allocations of risk may invalidate cost comparisons

In relation to the "not for profit" company "Glas", Ofwat noted that *"There may be practical difficulties for Ofwat in making robust cost comparisons due to the inclusion of profit and risk margins in contract prices"*.²⁶ However, the problem is more general than this specific example and potentially applies to any instance where the allocation of profit and risk within a business, or between in-house and outsourced activities differs from company to company.

Estimated efficiency differences do not yield an estimate of "X"

"Efficiency" scores are not sufficient to determine an X-factor for use in an RPI-X price control since the period over which assumed inefficiency is eliminated is arbitrary ("X" may also include an expected rate of productivity improvement for the "frontier" company *relative* to the economy as a whole).²⁷

**The Appropriate
Role of Yardstick
Methods**

There are a wide range of applications of yardstick methods, and a range of parties (including regulators) who may apply yardstick methods.

On the basis of our analysis we conclude that yardstick competition, which is reliant on unexplained differences in costs as a basis for setting allowed revenues, is inherently flawed and would not be expected to improve incentives. However, comparisons of various kinds are potentially informative to management and/or regulators. The question is how to define an appropriate role for yardstick methods that is consistent with promoting efficiency and customers' interests. We consider two approaches to deciding the appropriate role of yardstick methods:

- Through regulatory institutions (due process and appeals mechanisms) that make it unlikely that evidence of unexplained cost differences would be accepted as *prima facie* evidence of inefficiency.
- By setting out a role for comparative methods that we consider consistent with sound incentive based regulation.

Converging on an Appropriate Role for Yardstick Methods via Due Process

Well-developed requirements in terms of regulatory due process including opportunities for appeal of regulatory decisions may ultimately determine the appropriate role of yardstick methods. In particular, methods of regulation such as yardstick competition where it is unlikely to be possible to demonstrate that decisions result from a reasoned analysis of available evidence are unlikely to survive close scrutiny.

In the US, where clear administrative procedural requirements are legislated, and a reasonable assurance of cost recovery is enshrined in

the Hope Gas (1945) decision, the use of modelled or benchmarked cost estimates as a basis for setting allowed revenues is not the norm. One exception was the use by the Federal Communications Commission of Total Service Long Run Incremental Cost (TSLRIC) estimates, based on a hypothetical efficient network, to set local exchange company interconnection charges during the 1990s. However, the US Court of Appeals for the 8th Circuit struck down the use of TSLRIC in *Iowa Utilities Board v. the FCC* on 18 July 2000, noting:

“Congress intended the rates to be ‘based on the cost of providing the interconnection or network element’, not on the cost some imaginary carrier would incur by providing the newest, most efficient, and least cost substitute for the actual item or element... Congress was dealing with reality, not fantasizing about what it might be.”(p.7)

Benchmark cost estimates have also been successfully challenged in Australia (in relation to electricity distribution price controls),²⁸ and are under appeal to the Australian Competition Tribunal by Telstra in relation to telecommunications access charges derived using a network access model. Regulatory and legal precedent may therefore define the role of yardstick methods in Australia over time.

In the UK due process requirements and appeals routes are arguably weaker than in the US or Australia,²⁹ and may not therefore limit or define the role of yardstick competition to the same extent. In the longer term a stable and predictable role for the application of yardstick methods to regulation in the UK may only come about via reform of due process requirements and appeals routes.

Deciding an Appropriate Role for Yardstick Methods

Management within companies, owners, and market analysts apply benchmarking techniques to identify and/or reward efficiency. For example, senior managements' pay may depend in part on other companies' performance. The scope of regulatory benchmarking should not therefore be thought of as defining or limiting the extent to which benchmarking is applied overall.

Under price cap regulation a basis for setting the “X” factor in “RPI-X” formulae is required. One approach is to use past industry productivity *trends* as a basis for setting expectations. This represents a form of benchmarking – though it differs fundamentally from yardstick competition in that there is no need to normalise cost *levels*, and the starting point for the application of the “X” factor is actual company specific costs.

Regulatory parameters that are not directly observable may need to be estimated using industry wide data, for example, the cost of equity finance. However, in estimating such parameters judgement should be applied within narrow and well-defined bounds using agreed methods. For example, the Capital Asset Pricing Model (CAPM) could form the basis of an agreed method - though judgement is still required in deciding the appropriate comparator companies and, in particular, in assessing the equity risk premium.

Various kinds of benchmarks may be used to identify expenditures or areas of service quality performance that justify further scrutiny (including actual or proposed expenditure or service quality). However, unexplained cost differences are not assumed to represent *prima facie* evidence of *inefficiency*. The use of benchmarking as a means of identifying areas for further scrutiny in a regulatory context is set out clearly in the following from the Office of the Regulator General in Victoria (Australia):³⁰

“The Office’s preference was to set common benchmarks within each category [CBD, urban, short and long rural feeders], but it recognises that, even within those categories, there are differences between the distributors’ networks which explain the differences in their reliability levels to a significant extent...”

There are some differences in the distributors’ benchmarks, however, which are not readily explained by differences in network architecture. In developing their targets, the Office expects the distributors to consider whether those differences indicate that there is scope for significant performance improvement in the near term.”

While benchmarking is a useful and potentially powerful regulatory tool, we do not find support on efficiency grounds for the regulatory use of cost level benchmarks as a basis for setting allowed revenues. Rather than using cost benchmarks to decide what costs *ought to be*, price caps should be allowed to work in a stable and predictable environment, thereby improving incentives for utilities to *reveal* efficient cost levels over time. Various types of benchmarking may nevertheless be appropriate in determining regulatory parameters which cannot be derived from company specific information, and in identifying areas that may justify further detailed scrutiny.

Appendix A: Theoretical Underpinnings of Yardstick Competition

In his 1985 seminal article Shleifer proposed a scheme where the price a regulated firm receives for its service depends on the costs of identical firms (or heterogeneous firms with *observable* differences).³¹ Shleifer coined the term “yardstick competition” to describe this scheme, and concluded:

“If firms are identical, or if heterogeneity is accounted for correctly and completely, the equilibrium outcome is efficient” (Page 326)

The (theoretical) advantage of yardstick competition is that firms would have an incentive to operate efficiently, since price is set independent of their individual conduct, and customers would capture efficiency savings (mimicking a “price taking” competitive market outcome). However, the assumptions behind yardstick competition deserve careful scrutiny.

Shleifer set out the following condition for yardstick competition to be efficient (in relation to firms that are assumed to be *identical*):

“It is essential for the regulator to commit himself not to pay attention to the 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 (or, alternatively, can prove in court that firms chose to be inefficient and that their practices were imprudent), cost reduction cannot be enforced.” (p. 323)

For firms that are *not identical* different conditions apply:

“if (25) [the estimated cost function] is the exact version of (24) [the underlying cost function], and if q [the set of observable exogenous characteristics] is the complete list of characteristics accounting for diversity (i.e., the R^2 of regression (25) equals one), then this regulatory scheme yields first-best cost levels as its outcome...” (Page 325).

Shleifer noted that when heterogeneity is not fully accounted “*these problems may render the transfer insufficient to cover cost-reduction expenditures, thereby requiring a subsidy to run the scheme without bankruptcies*” (Page 325). The theoretical efficiency result for yardstick competition therefore depends on one or the other of two strong assumptions, namely:

- A willingness to bankrupt a potentially efficient firm - thereby harming efficiency; or
- Using subsidies to avoid bankruptcy – thereby undermining incentives.

Shleifer did not explore the implications for incentives and efficiency of each of these alternatives in his paper. However, either possibility raises serious doubts about the scope for yardstick competition to deliver improved regulatory outcomes in practice for the following reasons:

- The regulator may be unwilling to bankrupt a firm on the basis of econometric or other analysis of what costs ought to be – in which

case firms anticipating this would not respond as predicted by the theory (the regulator may effectively be required by law to avoid bankrupting suppliers).

- A full analysis of yardstick competition would consider the costs and benefits of bankrupting a potentially efficient firm in the final stage of the game. This wider game may not have an equilibrium.³²
- “*requiring a subsidy to run the scheme without bankruptcies*” amounts to bringing allowed revenues back into line with actual costs via a subsidy (or price adjustment), thereby reducing or eliminating the claimed incentive properties of the scheme since revenues would no longer be based on an exogenous benchmark.

Endnotes

- ¹ Kay, John (1993) "Efficiency and Private Capital in the Provision of Infrastructure" in "Infrastructure Policies for the 1990s" OECD, p.71.
- ² Jeff D Makhholm (14 November 1999). "Benchmarking, Rate Cases and Regulatory Commitment". Australian Competition & Consumer Commission Incentive Regulation and Overseas Developments Conference.
- ³ Littlechild, S.C. (25 January 1986) "Economic regulation of Privatised Water Authorities – A report submitted to the Department of the Environment", paragraph 1.18.
- ⁴ Makhholm, J and Quinn, M. (December 1997) "'X' Marks the Spot" *Public Utilities Fortnightly*.
- ⁵ Kahn, A. E. (1998) "Letting Go: Deregulating the Process of Deregulation", MSU Public Utilities Papers, The Institute of Public Utilities and Network Industries, p.92.
- ⁶ Weisman, D. L. (2000) "The (In)efficiency of the 'Efficient-Firm' Cost Standard" *The Antitrust Bulletin*, p.210.
- ⁷ The endogeneity for the "frontier" company arises since it has a 100 per cent weight in determining the position of the hypothetical efficiency frontier (this problem does not arise under Shleifer's original approach where a simple average or mean OLS estimate is used as the basis for setting allowed revenues). For Data Envelope Analysis the problem of endogeneity may be greater since more than one company may be on the estimated frontier.
- ⁸ Brian Williamson (May 2001) "Competition, policy stability and growth" in "Economic Growth & Government Policy" published by HM Treasury. Available at http://www.hm-treasury.gov.uk/docs/2001/growth_sem/sem_8.pdf
- ⁹ Moody's Investor Service (February 2001) "The UK Water Sector: Moody's Approach to Rating Highly-Leveraged Structures for Asset Ownership", p.6.
- ¹⁰ Op cit, paragraph 1.15.
- ¹¹ Fels, A. (March 1997) "Recent developments in competition policy – the impact on the Water and Wastewater industries." Speech to the Australian Water and Wastewater Association: 17th Federation Convention, p.9.
- ¹² A number of problems with the Ofwat approach have been documented in the past, for example, see Siôn Jones (October 1999), "Comparatively Poor? A Comment on the Ofwat and Ofgem Approaches to the Assessment of Relative Efficiencies", NERA Topic 22. Available at <http://www.nera.com>.
- ¹³ The latter is due, in part, to the limited number of observations available (the ability to discern a statistically significant influence falls rapidly as the number of explanatory factors increases relative to the number of cost observations).
- ¹⁴ Davidson, J. (1999) "OFWAT Efficiency Assessments using Econometric Models: a Comment". Available at <http://www.hyder.com>.
- ¹⁵ If any omitted variables are correlated with included cost drivers, then Ordinary Least Squares (OLS) estimates for included cost drivers will be biased. In addition, omitted variables bias means that statistical tests of the estimated relationship are invalid.
- ¹⁶ Stewart M. (December 1993) "Modelling Water Costs 1992-92 – further research into the impact of operating conditions on company costs" Ofwat Research Paper Number 2, p.6.
- ¹⁷ Ofwat (December 2000) "Water and Sewerage Service Unit Costs and Relative Efficiency: 1999-2000 Report". See Table 16 (p.41) for a summary of the most significant factors Ofwat allow for.
- ¹⁸ Dwr Cymru – Welsh Water (September 1999) "Ofwat's Efficiency Analysis – standing the Test of Time?" Available at www.hyder.com/dcww.
- ¹⁹ Driffield, N. and Davidson, J. (June 1999) "Econometric Analysis of Water Distribution costs", p.6. Available at: <http://www.hyder.com>.
- ²⁰ Op. cit. p.1.
- ²¹ One technique, known as stochastic frontier analysis, involves an *assumed* split of the error term (unexplained cost differences) in terms of random errors and a component due to inefficiency (with the distribution of the latter treated as asymmetric). Stochastic frontier analysis does not, therefore, overcome the problem of deciding whether unexplained cost differences are due to inefficiency or not.
- ²² Ofwat (1998) "Assessing the scope for Future Improvements in Water Company Efficiency – A Technical Paper", p.14.
- ²³ Koenker, R. and Bassett, Jr., G. (January 1978) "Regression Quantiles" *Econometrica*, Vol 46(1).
- ²⁴ Yatchew, Adonis (2001) "Incentive Regulation of Distributing Utilities Using Yardstick Competition", *The Electricity Journal*, January- February, p.57.

-
- ²⁵ Evidence of this has been found in panel data studies of US and UK manufacturing. See, for example, Foster, L., Haltiwanger, J. and Krizan, C.J. (November 1998) “Aggregate Productivity Growth: Lessons from Microeconomic Evidence” NBER Working Paper 6803.
- ²⁶ Ofwat (January 2001) “The proposed acquisition of Dwr Cymru Cyfngedig by Glas Cymru Cyfngedig – A position paper by Ofwat”, p.8.
- ²⁷ Relative productivity growth is estimated since the Retail Price Index (RPI) in the “RPI – X” formula already captures economy wide productivity growth.
- ²⁸ Office of the Regulator General (1 December 2000) “2001 electricity distribution price review – re-determination”.
- ²⁹ Brian Williamson (February 2001) “UK ‘Incentive’ Regulation – International Best Practice?”, in “CRI Regulatory Review 2000/2001 - Millennium edition”.
- ³⁰ Office of the Regulator General (June 1999) “Guidelines for the Preparation of Price Review submissions – Appendix A: Base-Case Levels of Service Quality”, p.5.
- ³¹ Andrei Shleifer (1985). “A theory of yardstick competition” *Rand Journal of Economics*. Vol 16(3), Autumn.
- ³² Shleifer demonstrates that for identical firms yardstick competition has a cost minimising Nash equilibrium (though only if the regulator can credibly commit to bankrupting an inefficient firm). If bankruptcy, or anticipated bankruptcy, is costly to efficiency the game involving identical firms may not have a sub-game perfect equilibrium (involving a Nash equilibrium in every stage of the game).

n/e/r/a

LONDON

*National Economic Research Associates
15 Stratford Place
London W1C 1BE
Telephone: (+44) 20 7659 8500
Fax: (+44) 20 7659 8501*

NEW YORK

*National Economic Research Associates Inc
50 Main Street
White Plains, New York 10606
Telephone: (+1) 914 448 4000
Fax: (+1) 914 448 4040*

OFFICES ALSO AT:

*Washington, DC • Los Angeles • Cambridge, MA • Philadelphia • San Francisco • New York
City, NY • Ithaca, NY • Chicago • Madrid • Sydney • Brussels*

© National Economic Research Associates
