Regulatory Policy 25th Anniversary Conference
Merton College, Oxford, 12th and 13th September 2016
The evolution of UK regulatory policy in retrospect: What has worked well? What hasn't worked well? What can be learned from the experiences?

Graham Shuttleworth[1]

Benchmarking: The Lessons for Regulatory Procedure (Speaking Notes)

Introduction

The ability of UK regulators to rely on benchmarking when setting revenues is not an innovation that is transferable to other countries. Rather, this reliance illustrates a persistent weakness in UK regulatory institutions:

- "Benchmarking" is the use of a statistical model or a linear program ("non-parametric" model) to set cost allowances for one regulated firm by identifying the costs of an efficient comparator firm. Benchmarking techniques were developed as an initial filter, identifying outliers that merit a closer look.
- However, the use of benchmarking to derive cost allowances at Regulatory Reviews has always been of dubious value, for at least three good reasons:[2]
  1. Setting revenues by the standard of "efficient costs" is not consistent with awarding the (risk-adjusted) cost of capital based on average returns in the stock market – which by definition only manages returns to firms of average efficiency. (Firms that manage to achieve efficient cost levels would offer investors higher than average returns.)
  2. Measuring "inefficiency" as the residual in a statistical (or non-parametric) model is intellectually dishonest. The residual may be due to any deficiency in the model, so assigning it to "inefficiency" is just prejudice.
  3. Regulators often convert a benchmarking measure of efficiency levels into a time trend (X-factor). In doing so, the regulator judges what is a reasonable *annual* rate of cost reduction. However, if the regulator already knows what that reasonable annual rate is, the calculation of an efficiency level by benchmarking serves no useful purpose.

- No agreed models have ever emerged from the practice of benchmarking, even though it is widely practised in the UK and elsewhere (e.g. Germany, the Netherlands, Austria, Spain, Belgium and recently Australia, among others). Each

---

[1] The author was for many years an expert at NERA Economic Consulting, as is now one of NERA's affiliated consultants. He has worked on regulation, and on benchmarking in particular, in many different countries. In preparing these notes, he benefited from the advice of erstwhile colleagues including Bill Baker, Richard Druce, James Grayburn and Jeff D. Makholm. However, the information and opinions stated in this note are solely the responsibility of the author.

[2] For further detail, see Shuttleworth, G. (2005), *Benchmarking of electricity networks: Practical problems with its use for regulation*, Utilities Policy, Volume 13, Issue 4, Pages 310-317, December 2005.

review sees the adoption of a different model – so the selection of input variables is highly subjective:
    - To illustrate the unreliability of model design, consider the latest review ("RIIO-ED1") of Electricity Distribution Networks (EDNs). Ofgem used benchmarking to decide which was the most efficient EDN to "fast-track" (to award a lighter form of scrutiny and a higher rate return). In the detailed investigation of "slow-tracked" EDNs, Ofgem revised the benchmarking model to the point where the fast-tracked DNO no longer appeared efficient.[3]
    - Ofgem has in any case been moving away from simplified statistical regressions with few variables, towards detailed comparisons of individual cost items (which at least offer the potential of like-for-like comparisons).
- To overcome these objections, German legislators enshrined the detailed method of benchmarking in law.[4] However, even they left the choice of input variables as a matter for regulatory discretion.
    - At one meeting attended by the author, when the consultant presented his analysis and the regulatory body (BNetzA) handed out the results, the officials charged with applying the rules looked very uncomfortable with the process, and with their responsibility for exercising discretion over the interpretation of the consultant's analysis.
- Looking at a wide variety of benchmarking exercises, it often appears as if the regulatory staff is happiest when the model awards most companies "efficiency" scores of 80-100% (which translates into X-factors of 0%-2% per annum), leaving only one company with a lower "efficiency" score and a reason to appeal.
    - The desire to achieve such an outcome, which gives a manageable level of appeals, probably feeds back into the regulator's choice of benchmarking model!

Why doesn't the appeals process lead to a more rigorous approach?

- UK regulators have never been willing to admit in public that benchmarking is a flawed process that creates risk. (This statement includes not just the frontline regulators such as Ofgem, Ofwat, Ofcom, etc, but also the bodies responsible for "referrals" on appeal, i.e. the MMC/CC/CMA.) The same is true of regulators in other countries. (Privately, regulatory officials often express doubts they would not acknowledge in print.)

[3] Grayburn, J., and Druce, R., *Reflections on the Successes of RIIO and the Scope for Future Improvement*, NERA Economic Consulting, Energy Regulation Insight, Issue 42, January 2016.

[4] Anreizregulierungsverordnung – ARegV, or Incentive Regulation Order, originally passed in October 2007; section §22, Sondervorschriften für den Effizienzvergleich (Special Provisions for Efficiency Comparison).

- One reason why benchmarking has never been tested, let alone rejected, as a regulatory technique is the focus of regulatory appeals in the UK.
    - The problem with benchmarking lies in the lack of objectivity over the selection of input data, model design and interpretation of results ("residual=inefficiency").
    - However, UK regulatory law accords wide discretion to regulators, similar to that accorded to government ministers, even though they are officials, and are not bound by political procedures and constraints.
    - Therefore, regulatory appeals in the UK have mostly focused on the *plausibility* of the regulator's analysis as a possible view, as judged by the officials of the appeal body. Regulatory decisions are not judged on their *objectivity.*
- There has never been a cross-examination of a benchmarking proponent in the UK capable of exposing the weaknesses, at least not of the sort that US institutions (statutory cost-basis, adversarial procedures) would have fostered.
- Such procedures would have shown benchmarking results to be subjective, selective and prejudicial (i.e. to rely on only one choice of model, when others are equally valid, and on the use of a residual to measure inefficiency, rather than the effect of other factors[5]).
- Unsurprisingly benchmarking is not used to form cost allowances in the United States as it would not withstand the more rigorous regulatory procedures applied there.
- It will probably also face major problems in Australia. An appeal against the use of benchmarking has already required the regulator to consider a wider range of models. A second appeal against this judgement is to be heard in October 2016.

The Institutional Flaw Underlying the Problem

- UK regulation still allows regulatory bodies to exercise wide discretion over the appropriate level of allowed revenues and over the methods used to set them. This discretion creates a high degree of regulatory risk which would only be overcome permanently by entrenching in law a standard for the level of revenues (similar to the Hope Gas 1944 decision of the US Supreme Court) and a standard for the administrative procedures to be used in reaching decisions (which is provided in the United States by the Administrative Procedures Act 1946).
- UK regulation still has no formal procedural standard that obliges regulators to provide evidence for every aspect of their decision – there is much more scope for using discretion or for giving credit to the regulators' expertise than in other countries.[6]

---

[5] Some techniques, such as stochastic frontier analysis, try to estimate the effect of unspecified "data errors" separately from the residual. However, that is just a redefinition of the residual. It is no less prejudicial to assume that the remaining residual can only be due to inefficiency.

[6] The requirements for a "procedural standard" are simple to discern from a document like the Administrative Procedures Act 1946, since they follow the simplest dictates of court procedure. Any regulatory body must: (1)

- For instance, when confronted by parameter estimates produced "in-house" (by their own staff or consultants) and different estimates provided by others, regulators may choose to rely entirely on the "in-house" estimates, even though their estimates have no greater status, and without saying why the conflicting evidence can be ignored. (This procedure is a bit like convicting a criminal because the prosecution has presented evidence he is guilty, and ignoring evidence of equal value that he is innocent – such as an alibi!)
- Until now, the weaknesses in regulatory decision-making have not really been challenged in the UK, since the appeal process is essentially a re-run of the regulatory review (in which the appeals body checks the *plausibility* of analysis, whilst exercising similar scope for discretion). There is no cross-examination of the original decision-makers to expose the weak basis of unreliable regulatory methods or decisions.

- The UK therefore has a weak "evolutionary engine", which permits the long-term survival of regulatory practices that are far from being the fittest (i.e. it is unable to eliminate any over-reliance on subjective and prejudicial methods like benchmarking).
- Generalising further, UK regulation has not been evolving towards good practices; instead, the UK has been subject to rapidly changing fashions and whims in regulation…only some of which would meet *international* standards of good regulation.

A Recent Change with Unforeseen Consequences (Perhaps)

- A recent change in the law has allowed regulated businesses and third parties to appeal against individual aspects of a determination, instead of rejecting the whole package. There are signs that this new procedure will fundamentally change how UK regulation works.
- After RIIO-ED1, two parties appealed to the Competition and Markets Authority (CMA): Northern Power Grid (as an EDN) and Centrica/British Gas (as a retail energy supplier)
- After certain grounds for appeal were winnowed out in the early stages of the process, these two parties were left appealing on six separate issues. Of these, the CMA accepted only two on appeal. However, the choices made by the CMA are significant.

provide an opportunity for the submission of all relevant evidence; (2) review all evidence submitted and explain in the decision why it is or is not to be taken into account in the decision; and (3) explain how the eventual decision is derived from the admissible evidence (noting any use of discretion when choices are inevitable). In practice, the regulatory body has to give equal weight to evidence of equal standing, or else it will be open to accusations of bias on appeal.

- The four cases that were rejected on appeal concerned particular *values*,[7] and the CMA chose not to challenge the regulator's discretion by imposing alternative values.
- In the two cases that were accepted on appeal, the CMA chose to reject the decisions of the regulator where they are *procedurally* flawed:
    - inconsistent with Ofgem's stated objectives – incentive scheme (IQI)
    - unsupported by evidence – Smart grid targets

- Challenging individual components of allowed revenue would in any case have been difficult to sustain without re-opening the whole review process.
    - The selection of individual aspects of a decision allows the appellant to "cherry pick" items where an appeal would work in its favour, whilst leaving any other items that are already favourable.
    - Overall, such an appeals process would tend to result in total package being too high or low overall (depending on whether the appeals came from a regulated firm or a third party).
    - The only way to check would be to re-run the whole regulatory review, as under the old "referral" procedure.
- However, by focusing on procedural flaws, the CMA is working to improve the *standard* of regulatory decision-making.
- By demanding a better link to evidence, the CMA will encourage regulators to make decisions more rigorously and more consistently, which should lead to better decisions, without the CMA itself having to provide alternative values or decisions.

This limited appeal procedure is relatively new and the CMA has not had time to develop anything like a standard procedure for such appeals. However, the recent decisions suggest that the CMA feels most able to accept appeals about procedural lapses in regulatory decision-making, rather than alternative values.

Perhaps, therefore, the UK is feeling its way towards a procedural standard for regulation – that will incidentally subject benchmarking to much closer scrutiny (and a more sceptical viewpoint) than previously.

It is only to be hoped that the new appeals procedure is allowed to persist for long enough, and to cover enough cases, for these lessons to be put beyond doubt.

[7] The four other issues were: Real Price Effects; regional labour cost adjustments; the cost of debt; and the incentive for reducing interruptions to supply.