

**A SCOPING EXERCISE ON QUANTITATIVE
ANALYSIS IN THE EVALUATION
OF EXTERNAL REVIEW OF
PUBLIC SERVICES**

PART 1: EXECUTIVE SUMMARY

A Report for HM Treasury

Prepared by NERA

21 May 2001
London

Project Team:
Michael Spackman
John Dodgson
Siôn Jones
Simeon Thornton
Jan Peter van der Veer

n/e/r/a

National Economic Research Associates
Economic Consultants

15 Stratford Place
London W1C 1BE
Tel: (+44) 20 7659 8500
Fax: (+44) 20 7659 8501
Web: <http://www.nera.com>

An MMC Company

TABLE OF CONTENTS

FOREWORD	I
1. INTRODUCTION	1
2. REVIEW BODIES AND REVIEW PROCESSES	2
3. REVIEW PROCESSES AS REGULATORY REGIMES	3
4. QUANTITATIVE ANALYTICAL TECHNIQUES	4
5. WHO MIGHT USE THESE QUANTITATIVE APPROACHES?	6
6. A PROGRAMME FOR FUTURE WORK	7

FOREWORD

NERA were invited by HM Treasury on 21 March 2001 to produce, by 21 May, a scoping report on the potential for quantitative analysis in the evaluation of public service review. The work is in support of a broader investigation by the Public Services Productivity Panel (PSPP), led by Sir Ian Byatt and Sir Michael Lyons, into the evaluation of public service review.

This document is Part 1 of our Report. Part 2 is a hard copy of a slide Presentation to the PSPP on 31 May. Part 3 is a set of "Notes to the Presentation", which develops a selection of specific issues.

Our study draws on inputs from a range of public service review bodies and other contributors. This includes formal meetings with the Heads of the four main review bodies identified in the terms of reference – the Audit Commission, HM Inspectorate of Constabulary (HMIC), Ofsted and the Social Services Inspectorate (SSI). We are most grateful to all, who gave most generously and constructively of their time. We are also impressed by the strong professionalism of all these bodies and the clear enthusiasm for continuous improvement – in their own performance as much as in the bodies under review.

We draw also on a wide range of less formal discussions with officials from these bodies, and from the Treasury, the Regulatory Impact Unit (RIU) Public Sector Team and the National Audit Office, and with academics and other interested parties; and on the discussion at the PSPP Follow Up Meeting with reviewers of 10 May.

We also draw on a range of written material, mainly from the websites of the inspectorates (together with material provided by them) and the Cabinet Office (RIU and Performance and Innovation Unit) and the *PublicNet* site. We have also benefited from the parallel work for the Treasury of the Centre for the Study of Regulated Industries (CRI). Although our conclusions differ from those of the CRI, we wholly endorse their background papers, in factual reporting and analysis, to the extent that we address the same issues.

In discussion with the reviewing bodies, and others, we have explained that we will not quote what they say in any attributable form. Where we record our understanding of the views of particular bodies, these understandings are in any case derived from several sources. They should not necessarily be attributed to those whom we met in formal meetings, nor taken as formal institutional positions.

1. INTRODUCTION

We approached this scoping study as a relatively conventional application of public service economic analysis. That is, as a task of establishing the objectives of external review bodies, examining how alternative ways of doing things might be identified, and drawing conclusions about what might be done to quantify and where possible value the comparative costs and benefits. However it became clear during our investigations that, to come up with useful recommendations, we needed in two respects to take a slightly broader perspective.

The first of these broader perspectives is the need to extend our field of interest from the review bodies alone to include the review process of which they are a part, and to find a way of describing this complex world in a systematic way. The second is the question of who might use the quantitative approaches which we identify.

In this Executive Summary we first address the issue of review processes and then set out a framework for describing what might be quantified. This is followed by a discussion first of quantitative techniques and then, in our view crucially, who might apply them. We conclude with suggestions for further work.

2. REVIEW BODIES AND REVIEW PROCESSES

For assessment of the public *service providers* the distinction between the body and the policy it is applying is often not important. National policies may have important impacts on outcomes – as with literacy hour on school outcomes and the aftermath of Macpherson on policing outcomes. But the management of a school or a police force can be assessed in terms of the level and change of educational or police-related outcomes (and similarly for a social services department, or a local authority Best Value Service against its Best Value targets).

The management of review bodies can also be assessed in some aspects with only limited reference to the wider policy scene. This applies to their internal management, for example policies towards staff, or IT, or budget planning. However the management performance of a review body cannot be assessed so readily in terms of the impact of its activities on public service outcomes, for a number of reasons.

- Review bodies typically have a range of functions, some extending well beyond review. For the impacts of most of these functions, such as policy advice, there is no useful counterfactual. The same applies to the effect on the bodies under review of the mere existence of a review process.
- In assessing a review body's work the *scale* of its activities is often a major concern, which cannot be isolated from structure or style – in contrast to a service provider, where the scale of activities is seen as a budgeting issue, distinct from whether it is doing a good job with the resources available. But the scale of the review activity is determined largely by other authorities.¹
- There is in any case much to be said for improvements in service delivery to be clearly 'owned' by the service provider, and not claimed in part by the review body, even where this may be feasible.

Broadening the “evaluation of review of public services” to include the review process also brings other advantages. It allows a clear, uniform structure for organising the description and examination of this complex policy field. And it fits well with the concepts of joined-up government and proportionate regulation.

¹ For example Ofsted now have as one of their SDA performance targets the percentage of schools whose performance has improved since the previous inspection. This is sensible. But this measure tells us more about the regime that Ofsted is operating, and perhaps about wider education policy, than about the efficiency with which Ofsted is applying it.

3. REVIEW PROCESSES AS REGULATORY REGIMES

As a starting point for understanding review processes we commend a framework for the analysis of regulatory regimes recently developed in the LSE.² Coherent discussion of regulation is very difficult without a framework of this kind.

This framework provides, in the authors' words, "a tool for mapping and describing regimes, to identify commonalities and differences; contributing to better debate about how and why regulation varies across public policy domains; and [of most relevance to this Treasury study] the assessment of institutional and policy design". The framework's basic features are:

- Disaggregation of the regulatory process into the three standard control theory components of: '*information gathering*', '*policy making*' or '*rule setting*', and '*behaviour modification*'.
- Distinguishing between regime Context and regime Content.
- Disaggregating Context and Content into elements (and successive levels of sub-elements, where this is helpful).
- Constructing a matrix for the regime, with the three control components as one dimension and the elements of Context and Content as the other.

The three elements of regime Context are: the *inherent characteristics of the regulated activity* (such as the incentive structure with no regulation; whether it is a local authority, NHS or central government service); the *public and media profile*; and *the extent and distribution of organised groups* (such as trade unions and consumer groups). Even across these elements there are important differences between public service regimes – for example in the public and media profiles of Ofsted (high), HMIC (low) and the SSI (low but potentially high in cases of failure) – which have an important bearing, for good and ill, on what the review bodies do and how they do it.

However quantitative analysis to assess institutional policy and design is more concerned with regime Content, of which the elements are defined as: *size* (e.g. costs; political aggression), *structure* (e.g. organisational complexity) and *style* (e.g. rule orientation; support versus 'name and shame').

² C Hood, H Rothstein and R Baldwin, *The Government of Risk: Understanding Risk Regulation Regimes*, Oxford University Press, 2001 (forthcoming).

4. QUANTITATIVE ANALYTICAL TECHNIQUES

Broad areas to which analysis can be applied, to examine the efficiency and proportionality of individual regimes are:

- internal efficiency of the reviewing bodies;
- comparison of costs and benefits of new regulatory proposals; and
- strategic analysis of existing regimes.

Comparative or coordinated examination across regimes can be applied to:

- regime elements; and
- the development of regime style and methodology.

Internal efficiency of the reviewing bodies is important and can always be improved, but incentives to achieve and demonstrate this are already strong. It does not appear to be a high priority for new development.

For assessing new regulatory proposals the concept of “Regulatory Impact Assessment” (RIA), including “Compliance Cost Assessment”, is well established for regulation of the private sector. Both these concepts, especially the latter, have important limitations, but both appear to be as applicable to the public sector as the private sector. The RIU however suggests that RIA is not applicable to the public sector and has proposed instead a more administratively focused tool³, which was discussed at an inter-departmental meeting on 2 May.

We discuss these issues in Part 3. We conclude that, as the RIU propose, there should be stronger requirements for proposed new public sector review regimes, and for new review bodies, to be rigorously appraised, and that the main issue is to establish incentives to bring this about. The techniques needed, whether or not they are given a special name in this context, are those of conventional public service policy analysis.

The term “cost-benefit analysis” also appears in current discussion of public service review. We feel this term is best reserved for its universal technical usage, which is to describe analysis which makes substantial use of valuations of non-marketed impacts, such as the valuation of leisure time, of risk reduction, or environmental impacts. There is little scope for this in the context of public service review.

³ The term “Impact Assessment Tool” is used by the RIU as a generic term to include “Regulatory Impact Assessment” for private sector regulation, and a related but different procedure for the public sector, for which the RIU have proposed the term “Regulatory Effects Framework”. We use this latter term for convenience, although Ministers have yet to approve this or some other term.

As examples of strategic analysis of specific regimes we understand that Ofsted is looking at the question of whether the present scale of school inspection is still appropriate, and the SSI is debating the question of whether the lighter touch approach (or risk-assessment approach in HMIC language) is appropriate for social services, where there is a risk of serious harm to individuals if a local authority department fails. There is wide scope for such analysis. This should however be a normal part of policy appraisal and evaluation.

We see considerable scope for the comparative or cooperative analysis of regime elements. An outstanding example, narrowly focused and effective, is the ‘paperwork’ studies by the RIU Public Sector Team – focusing on the paperwork load of schools, police forces and GPs. We believe that systematic examination of these regimes, using the framework described above, would yield a rich menu of issues. This analysis would be useful to central departments, policy departments, the reviewing bodies and the service deliverers under review. It might include comparative analysis of for example: the use made, by whom, of information obtained by review; the effectiveness of surveys to measure the value of review; and the procedures (who and how) used for adapting the scale of review over time.

Perhaps richest of all, in terms of potential for new analysis, is the scope for comparative or cooperative development of review style and methodology. The public service review regimes are striking in both their differences and their similarities. Many of the issues which they face are the same, but since the regime contexts are different the best approach varies – both between and within regimes. Potential issues include the categorisation of bodies under review – for example to prioritising between the “unambitious average” and the “laggards”; improving early warning indicators of the need (or otherwise) for review; the continuing development of performance measurement; handling of the media; the role of information for the public and how best to present it; the balance between support and sanctions; the extent to which information requirements can be simplified (for example by combining, with an acceptable loss of detail, requirements which are similar but not identical); or the use of external advice on anything from document design to the quality of the review body’s reports to the conduct of surveys.

We could conceive of potential uses for almost any analytical technique, including sophisticated cost benefit analysis, econometrics, benchmarking, multi-criteria analysis, and survey analysis. It seems clear that the scope for more sophisticated techniques requiring high quality data is limited. But there is ample scope for analytically simpler, but conceptually and administratively no less demanding approaches to data collection and interpretation.

There may be scope for commissioning more independent analysis. Enquiries by reviewers into their own impact, for example, are unlikely to provide an accurate picture.

A prior requirement for the development of review style and methodology is a technical and professional framework within which the questions which need analysis can be asked and addressed in a coherent way.

5. WHO MIGHT USE THESE QUANTITATIVE APPROACHES?

We noted above that there is considerable scope for quantitative analysis of *single regimes*, including the analysis of proposed new regulations and, especially of new review bodies. These requirements have often developed as political “imperatives” and appear not to have been examined as closely as the taxpayer or public service user might hope. This is an issue which probably merits a higher profile in the Spending Review process (as we briefly discuss in Part 3) and in central guidance.⁴

The review bodies may often have a major promotional role in this work and do much of the analysis, but the prime responsibility for seeing that such analysis is developed lies with the policy departments.

More complex is the development of comparative and cooperative analysis. The examination of specific elements can be promoted and coordinated by the RIU, as well demonstrated by their studies of paperwork. However no central department is equipped to coordinate work on the development of review style and methodology.

We found a consistent interest by reviewing bodies in work of this kind, combined with considerable caution, among the professional reviewers, about the dangers of construction of a new regime to “review the reviewers” which would reduce their effectiveness.

This is moving to the edge of and perhaps beyond our terms of reference, but it appears to us to be so central to the development of comparative and cooperative analysis in this growing and important field that we are bound to stress it. We see little prospect of substantial development without some cooperative forum, supported by the review bodies, which goes beyond the various present structures such Heads of Audit and the Best Value Forum. We discuss these issues in Part 3.

⁴ These concerns continue to carry much higher weight in the regulation of the private sector than in the regulation of the public sector’s own activities – presumably because, in the “regime context” of private sector regulation, the influence on Government opinion of organised producer groups is stronger, and the media often oppose regulation (whereas in the public sector they press for ever more).

6. A PROGRAMME FOR FUTURE WORK

There is scope for more analysis of specific regimes, but this is a matter of applying more firmly the established techniques of policy analysis and review, as set out for example in the Treasury green book and the RIU guidance on Regulatory Impact Assessment. This analysis appears not to have been applied rigorously because proposals for increasing review have been seen as political imperatives. This is an issue for the central departments to address, prompted and steered by the Treasury, perhaps within the framework of the Spending Review.

The analysis of *new* proposals is at present being promoted by the RIU. This initiative deserves support, but the proposed Regulatory Effects Framework appears to have technical limitations which it would be helpful for the Treasury to address.

There is also considerable scope for comparative and cooperative analysis of public sector review.

The promotion and coordination of cross-cutting microanalysis of specific elements of these regimes can be well handled by the RIU, using analytically simple (albeit administratively demanding) techniques.

However the development of regulatory style and methodology, upon which the effective appraisal and evaluation of regimes largely depends will only take off with the evolution of a collaborative forum, probably with a small permanent secretariat, committed to developing this work. If such a procedural development were in prospect there is scope for useful work to fill out and prioritise the potential studies which might be undertaken.