

Measuring Quality as Part of Public Service Output

Strategy following consultation

Foreword

Measuring the output and productivity of our public services, such as health and education, is important for public accountability. Public services are around a fifth of total Gross Domestic Product. Taxpayers, users and providers of public services have an interest in how government spends our money on these services and in whether services are good value for money. This paper is concerned with techniques for measuring changing levels of public service output, and productivity, over time, with particular emphasis on measurement of the quality of services.

Independent, authoritative and reliable information and methodologies are needed, in an area of political and public interest. So far as possible, methods should be agreed internationally and enable cross-national comparisons of productivity and productivity change. The ONS UK Centre for the Measurement of Government Activity (UKCeMGA) was set up in 2005 to develop these methods, following the Atkinson Report on *Measurement of Government Output and Productivity for the National Accounts*. UKCeMGA feeds information into the National Accounts and publishes public service productivity articles. Information used for productivity articles may differ from the National Accounts, where this can be justified, in particular by using more 'developmental' methods.

For transparency and credibility, developing methods in this area should be subject to public debate. UKCeMGA opened a consultation in September 2006 on some key methodological issues, including their application to developments on measurement of education and health services. Information on views expressed during the consultation has already been published. This document sets out the ONS conclusions and forward strategy, as a result of consultation and further internal development.

The strategy sets out some key concepts and a development path. Expert engagement in developments will continue to be important. ONS will continue to consult, with documents made available on its website and discussion encouraged by various means. This includes partnership with government departments with expert knowledge of different services, and use of academic research commissioned by ONS or by departments. The independent Advisory Board for UKCeMGA will continue to ensure that work priorities and decisions on methodologies have objective credibility.

We thank all involved in the consultation.

KAREN DUNNELL
National Statistician

PETER SMITH
Chairman, UKCeMGA Advisory Board

Summary

This paper sets out:

- the overall strategy for measurement of public service output, with particular reference to quality;
- the development programme needed to take the strategy forward;
- specific decisions on measurement methods for health and education.

The strategy follows from the Atkinson Report on *Measurement of Government Output and Productivity for the National Accounts*, which advised that output should be measured in a way which takes account of quality change. The strategy takes account of views expressed in consultation over the last year. It proposes a **conceptual framework** for the measurement of quality of public services. The framework will allow credible, transparent estimates of public service output which use the best data sources and research methodologies available, and make informed estimates when data do not allow definitive conclusions. A key objective is that estimates reduce the errors of measurement associated with not measuring the quality of output at all.

The **key concepts** are:

- defining the right *unit of output*;
- *dimensions of quality* – to be defined for each service within an overall framework of
 - The extent to which the service succeeds in delivering intended outcomes;
 - The extent to which the service is responsive to users' needs.
- *techniques* – for example, measuring the actual contribution the output makes to outcomes;
- *capacity for benefit* – variables affecting the way individual users can actually benefit from the output delivered.

The strategy identifies work needed on three **methodological issues**:

- *the relative importance of quality dimensions* – needed for research, or informed judgement, on how much weight should be given to different aspects of quality;
- *using value weights instead of costs weights* – cost weights now used to measure output may not reflect the value of public services to users, but will continue as a pragmatic solution unless value weights can be developed systematically;
- *time lags* – techniques are needed to recognise that the effects outputs have on outcomes may be delayed or spread over a number of years.

Transparency: the paper sets out the ways in which ONS intend to ensure the relevance and robustness of particular concepts, techniques and data sources used in this work.

The paper concludes that the impact of public services on the market sector (**complementarity**) is relevant for development of a new measure of welfare, but will not be used to adjust measures of public service output in the National Accounts or productivity articles.

The strategy for measuring quality as part of **education** output is:

- a more refined use of GCSE results;
- examination of measures using Key Stage results to measure progress within school years;
- developments to take account of diplomas and A levels;
- developments to measure quality of under five pre-schools education;
- all to be set in the context of an extended set of quality domains based on the outcomes covered by the 'Every Child Matters' framework;
- no upward adjustment to education output to take account of complementarity between private and public sectors.

The strategy for measuring quality as part of **healthcare** output is:

- quality dimensions to put more weight on health gain than patient experience;
- more work on measurement of health prevention as part of health output;
- desirability of systematic information on health gain from treatment;
- continued work on value weights for healthcare, but productivity articles will not use a mixed 'value and cost' weight combination;
- no upward adjustment to health output to take account of complementarity between private and public sectors.

There will be **further work** to develop the quality measurement framework through research funded by the Invest to Save Budget. This has particular focus on adult social care and under fives education, and tools for local use to measure the quality of services, provided by public, private or third sector providers. The project will also be used to strengthen methodologies for quality measurement which can be used in other areas of public service.

1. Aims and Objectives

- 1.1. This paper ends a formal consultation programme (see section 2) and sets out a forward strategy. The paper sets a forward direction for each of the main propositions discussed in the consultation papers, but does not give them all equal coverage. It introduces some additional ideas based on reflection on discussion at the consultation seminars.
- 1.2. The objectives of this paper are to set out:
 - i. the overall strategy for measurement of public service output, with particular reference to quality;
 - ii. the development programme needed to take the strategy forward;
 - iii. specific decisions on the future of various measurement methods discussed in the health and education papers.
- 1.3. The paper is meant to be freestanding and avoids technical details, but the issues can be understood in more depth by reference to the consultation papers, public service productivity articles, the Atkinson Report and other material on the UKCeMGA website (see Annex A).
- 1.4. The rest of the paper is set out as follows:
 - section 2** provides background to the consultation process;
 - section 3** develops a robust framework for measuring quality as part of the output of public services;
 - section 4** sets out conclusions on private and public sector complementarity;
 - section 5** sets out plans for measurement of quality as part of education output;
 - section 6** sets out plans for measurement of quality as part of healthcare output;
 - section 7** sets out plans for measuring output of other public services e.g. social protection and public order and safety;
 - section 8** sets out the next steps.

Annexes include information on reference material, the Cabinet Office Guidelines on consultation, how to find more detailed information on the consultation responses and seminar summaries, and the report of the expert economist panel on complementarity.

2. Background to Consultation

- 2.1. This section explains the history of ONS work on output and productivity of public services and the context for the recent consultation.

Context: Public Accountability

- 2.2. Public services are services provided to users which are provided by a public body, or purchased by a public body for use which is free, or has only nominal charges, at the point of delivery. They do not include cash transfers (e.g. state pensions) but do include the administration of such benefits. The amount of tax-payers' money spent on public services can be identified relatively easily. The work covered by this paper aims to measure what is provided by the different services, in terms other than 'what was spent'. Other traditional measures like 'number of teachers' or 'number of hospital beds' also describe *inputs* to services, not the *outputs*.
- 2.3. For many purposes of public accountability, the main interest is in *outcomes* for the public rather than output. For example, a reduction in reported crime, and fear of crime, is an outcome for citizens. It may be achieved by more police, or new policing techniques, or by other changes like better security devices on cars. Provision of public service is seldom the only way of achieving desired outcomes, and a proper role of government is to promote action by other parts of society. Nevertheless, the amount of Gross Domestic Product (GDP) spent on public services is high, and it is important to be able to measure whether funds are being used successfully. For criminal justice, we seek to measure the *output* of the police, courts and offender management system, in a way which is linked to their impact on outcomes.
- 2.4. Productivity measures compare the volume of outputs with the volume of inputs (after taking out pay and price changes). The techniques discussed in this paper are about change in productivity over time, and are used for 'all England' or ideally 'all UK' output. Some other work on productivity measurement focuses on the productivity of different providers – for example, comparing schools, or comparing the efficiency of large and small hospitals. These are also important, particularly in helping service providers realise how they can improve. But they do not directly answer the question of 'are we getting more or less, over time, from publicly funded services, relative to what we put in?'

Context: National Accounts

- 2.5. Public services are included in the National Accounts, which measure economic activity overall, alongside the market sector (e.g. financial services, manufacturing industry). Up to 1998 the ONS, like other national statistics offices, measured the output of public services for the National Accounts by assuming the outputs were equal to the inputs used to produce them. This meant that output could only ever grow at the same rate as inputs and measured productivity growth would always be zero (since productivity is calculated as the ratio of output to inputs).
- 2.6. From 1998, and following European directives, the ONS started to measure public service output by direct methods and made good progress. However, the data and methods available to measure output directly were far from ideal and progress was rather piecemeal. Recognising the importance of moving this work to a higher standard, the then National Statistician, Len Cook, asked Sir Tony Atkinson to carry

out a review. The Atkinson Report was published in January 2005 - '*Measurement of Government Output and Productivity in the National Accounts*' - and set out principles and recommendations for moving this work forward.

2.7. Three of the principles in the Atkinson Report are particularly important for this paper:

- **Principle A:** the measurement of government non-market output should, as far as possible, follow a procedure parallel to that adopted in National Accounts for market output.
- **Principle B:** the output of the government sector should in principle be measured in a way that is adjusted for quality, taking account of the attributable incremental contribution of the service to the outcome.
- **Principle C:** account should be taken of the complementarity between public and private output, allowing for the increased real value of public services in an economy with rising real GDP.

2.8. The terms and issues are explained further in the rest of the paper – but it is important to note that Principle B links output back to outcomes, though only to the extent that services actually impact on outcomes. The full list of Atkinson Principles is at Annex C.

Developments Since the Atkinson Report

2.9. In July 2005, the National Statistician set up a new directorate in the ONS, the UK Centre for Measurement of Government Activity (UKCeMGA), to take forward implementation of the Atkinson principles and recommendations. Len Cook also gave an undertaking that any proposals for significant changes to the methodologies used to measure public service output and productivity would be the subject of wider debate and public consultation, before being considered for adoption in the National Accounts.

2.10. Since its launch UKCeMGA has made substantial progress and has published a series of Public Service Productivity articles on health, education, adult social care and social security administration, and a paper on improvements made to central government expenditure data. The Department for Education and Skills (DfES), the Department of Health (DH) and the Department for Work and Pensions (DWP) published articles on measurement of output of particular services, including proposals on quality adjustment.

2.11. In summer 2006 ONS launched a structured consultation on some of the cross-cutting methodological issues on quality measurement as part of output, and on complementarity of public and private services. This included consultation on specific issues on quality measurement for education and health. The aim was to ensure the work benefits from expert opinion and proceeds on a widely supported and well-based path. This was to be achieved by facilitating understanding and discussion of current methods and the underlying issues.

2.12. Specific success criteria included:

- that the results provide a clear steer for the next phase of development of the

methodology, and allow an assessment of the extent to which a consensus exists regarding the specific methodological issues that are posed;

- the consultation has benefited to the maximum extent possible from relevant expert opinion, both practitioners and analysts; and
- that everyone who would wish to do so has had the opportunity to contribute to the debate.

2.13. The consultation programme was carried out in three phases:

- Establishing the Principles** - the first consultation ran from 18 September to 11 December 2006 and focused on two key methodological questions common to the measurement of all public service output:
 - **Adjusting output for quality change** - to discuss the principles that should underpin the measurement of quality change in the public services. (Principle B of the Atkinson Report);
 - **The value of public services in a growing economy** - to discuss the economic foundations for recognising public and private sector complementarity in the measurement of output of public services. (Principle C in the Atkinson Report).
- Measuring the output of the Education Service** - the second consultation ran from 2 November 2006 to 22 February 2007 and focused on how to improve our current methods for measuring education output and quality.
- Measuring the output of the Health Service** - the third and final consultation in the programme ran from 12 December 2006 to 3 April 2007 and focused on how to improve our current methods for measuring healthcare output and quality.

2.14. Each phase began with publication of a consultation paper, published on the ONS website and sent to bodies expected to take an interest. Seminars were held to enable deeper discussion, as well as inviting written comments. The written responses to the consultation and the summary of the seminars were published on UKCeMGA's web pages (www.nationalstatistics.gov.uk/ukcemga) within six weeks after the end of each consultation.

2.15. The consultation was conducted in line with the Cabinet Office Guidelines, set out in Annex B.

2.16. This paper forms the formal response to the consultation programme and sets out the strategy that UKCeMGA will adopt in order to take this work forward.

3. Conceptual Framework for Including Quality in Measures of Public Service Output

Introduction

- 3.1. The consultation set out some proposals and asked questions about methods of quality measurement in general, in the context of output and productivity measurement, with more specific proposals for school education and health. The first consultation paper, *Establishing the Principles*, emphasised the need for a framework for quality measurement, to be used consistently for different services. This strategy paper begins with further explanation of the conceptual framework to be used. This helps to understand different types of error and the way they may be reduced, by research or (in the interim) by justifiable approximation. It sets out the 'difficult questions' where further technical work is needed to establish practical guidance on acceptable techniques: ONS will take this forward, in discussion with others, and will regularly publish work in progress and seek comments from those with an interest.
- 3.2. One theme of the consultation was whether different measures should be used for different purposes, and there was agreement that this was so. In particular, methods used for the National Accounts should adhere to international guidance, and there should continue to be a high standard for work leading to revisions of the National Accounts. There was agreement too that developmental work on quality issues should continue and that a value based approach (see paragraphs 3.22 – 3.24) should also be pursued, with regular publication, to inform understanding of complex issues and share progress with other interested agencies and users. The results of such work, even if they were still experimental, could appropriately be presented alongside established series, for example in productivity articles.
- 3.3. Later sections of this paper discuss particular measurement approaches for education and health based on some data which are or might be available. The techniques set out for further development and use should not be regarded as giving 'perfect' measures. They are put forward because there is sufficient reason to judge that they reduce the error caused by ignoring quality altogether. Their use and interpretation should be based on that understanding.
- 3.4. The issues discussed in this strategy are also part of a Quality Measurement Framework project, for which HM Treasury has provided specific funding. The links are discussed further in paragraphs 3.33-3.35.

Why Quality?

- 3.5. The consultation paper, *Establishing the Principles*, explained the context for measuring public service output and the reasons why the Atkinson Report, and international guidance, regard quality as an intrinsic part of output. Output measures are generally constructed from measures of the volume of different types of 'product' and their price or unit cost. In the market, quality is an intrinsic part of the concept of a product – a luxury holiday, or a basic one, priced accordingly. For public services, it can be harder to distinguish relevant aspects of quality, as distinct from the quantity of a service (number of operations or number of school pupils), but measures will be misleading if quality issues are ignored. UKCeMGA's work on productivity analysis is particularly focused on aspects of quality which may change

over time, and techniques for incorporating measures of quality into time series. While there is much interest in measuring aspects of quality in different public services, and improving quality for users, there is much less experience of constructing a single measure of output which combines all aspects of quality, and volume and cost, for a service. Without such a measure, it is not possible to say how value for money has changed over time.

- 3.6. Where public services are delivered to individual users, measurement of quality as part of output is akin to considering the way quality affects price. We may not know to what extent different aspects of a luxury holiday contribute to the price (comfort of bedroom fittings? Amenities in hotel and grounds? Free taxi from airport? Range of choice at breakfast?), but we recognise that individuals get 'utility' from these features and make a choice.

Quality Measurement: Key Concepts

- 3.7. This section sets out some key concepts on quality measurement. Some were explained in the consultation papers; others were raised in discussion at the seminars.

Defining the Right 'Unit of Output'

- 3.8. Most of the current measures used for public service outputs are quantities of 'activities' – for example, weeks of nursing home care, GP consultations, a pupil attendance at school for one day. These tell us something has happened, but not whether it was successful in achieving its intended purpose, or whether the user regarded the experience as satisfactory. The ideal way forward would be to define a unit of output in a way which included quality – for example, 'a unit of healthcare which successfully diagnoses and treats an illness' – this might be made up of a long sequence of separate consultations, tests and treatments. Discussions in consultation seminars showed some disquiet with the relevance of some of the current quantity measures, though data limitations were recognised.
- 3.9. It is very important that the quantity measures are chosen suitably. In addition, much more development is needed on quality issues. It is possible to take an index which measures examination attainment, and combine it with a quantity index based on school attendances to obtain a quality adjusted volume measure of output. But there is a hidden issue, of whether '1 per cent more success in exams' is equal to '1 per cent more pupils at school'. If the ideal unit of output is 'pupils who attend school regularly and pass exams successfully', then the two types of increase seem to be equivalent. If the ideal unit of output is more like 'pupils who attend school, avoid all the risks and problems of truancy, develop social skills and self confidence, and a range of competences which give them good life chances in future', then there is more question about the exact technique for matching the examination improvement to the quantity change.
- 3.10. There is no easy resolution to difficulties of finding the ideal unit of output, for each service. But it is always useful to ask 'what are we really trying to measure?', and then consider how far the available data and techniques approximate to that.

Dimensions of Quality

- 3.11. The consultation paper, *Establishing the Principles*, proposed that the first step in

measuring quality was to establish a 'robust and complete set of quality dimensions'. These should be based on research, international collaboration or other steps to ensure credibility. Quality dimensions, and the precise vocabulary used to describe them, are likely to differ between services. However, there are likely to be similarities, with quality dimensions falling into two broad areas:

- The extent to which the service succeeds in delivering intended outcomes;
- The extent to which the service is responsive to users' needs.

3.12. The second dimension can encompass factors like provision of clear information about services available, the experience of users, respect for dignity of users, and convenience of the way services are provided. These may be regarded as less important than tests of 'delivering intended outcome', but can be measured independently and are known to be important to users. 'Equity' is not included in this list, though it features in other performance frameworks, and is discussed below in paragraphs 3.28-3.32.

3.13. A 'meta framework' for quality would be helpful, if it could be agreed to be relevant to a wide range of services, to help test the validity of proposed dimensions of quality in particular areas. More exploration is needed of whether the two part division is helpful in areas where users are obliged to accept a service (e.g. aspects of criminal justice) and where services are not primarily directed at individual users (e.g. fire prevention). ONS will continue to work with experts, including international bodies (for example, the OECD) to agree the most suitable set of quality dimensions for each public service. This will be linked to separate work within ONS, and across Whitehall, to develop a framework of measurement for societal wellbeing.

3.14. There was some discussion of whether the aim should be to identify a 'complete' set of quality dimensions, or 'representative' quality dimensions. UKCeMGA considers the aim should be a 'complete' understanding of what aspects of quality matter to users, even if they cannot all in practice be measured. This helps to reduce the risk that quality indices are developed on the basis only of what can easily be measured, ignoring areas which can be understood in conceptual terms but not measured.

Services to People: Capacity for Benefit

3.15. The characteristics of service users vary, at any time. For example, schools include pupils with special needs, whose ability to benefit from education may depend on provision of extra help, at additional cost. The proportion of pupils with special needs may change over time, or techniques for helping address needs may change. The results of services, and their unit cost, depend on the characteristics of users, as well as the skills and resources of service providers. If these 'capacity for benefit' factors are ignored, there may be error in measuring output and productivity of public services. We need to measure whether the service delivers 'the best attainable outcome for the individual user' rather than regarding 'intended outcome' as an absolute standard. This concept has been set out in research commissioned by DH from the Personal Social Services Research Unit with reference to measuring output of adult social care services, but it applies more widely.

Techniques of Quality Measurement

3.16. The consultation paper summarised the explanations in the Atkinson Review on approaches to measuring quality. The first approach is differentiation: to ensure

different types of service are measured separately, as part of the difference may be in the quality of what is provided (as with the holiday example). ONS has been extending the detail and breadth of coverage of different areas of output and will continue to do so. The consultation noted that the aim was not so much 'a large number of separate activities' as to ensure that different types of output are homogenous. As an example, counts of hospital activity distinguish healthcare resource groups which are based on diagnosis, procedure, and resources used. While differentiation is important to ensure the quantity of services is measured well, it cannot capture all aspects of quality.

- 3.17. The other two Atkinson recommendations on quality measurement were to use quality indicators which show the successful delivery of outputs, or alternatively to use evidence on change in outcomes which can be attributed to the incremental contribution made by public services.

Indicators of the successful delivery of outputs (e.g. survival rate after operation): these need to be tested carefully against the conceptual framework of quality dimensions to see if available data measure what is actually important to users, and whether one or more indicators can be found that are sufficiently representative of a range of experience. For example, hospital survival rates could be regarded as a useful indicator of quality of emergency appendicitis operations, but not as a measure of quality of dermatology services. Other technical issues also need to be considered: combining evidence from several separate indicators which measure the same dimension of quality (or choosing between them), handling breaks in time series as definitions change; whether indicators are constrained by a floor or ceiling effect (once 'accuracy' reaches 99.9 per cent, there is not much chance of further improvement); whether indicators are distorted in some way by 'better measurement' or 'better at doing exactly what is measured' rather than 'better delivery for users'. Quality measures based on user surveys may be helpful for some quality domains, but there are issues on whether 'subjective' measures from successive sample surveys can be used as a time series – they may be distorted by changing expectations. As noted in paragraphs 3.8 – 3.10, the approach to combining quality measures with quantity measures needs care, judged against an ideal 'unit of output'. ONS will work with experts to develop further practical guidance on acceptable techniques in these areas, aiming to understand and reduce error rather than to set a standard for 'perfect' measures which may be unattainable.

Measuring quality of output through impact on outcomes: this is a good technique where it is possible to be sure that the change in outcome is entirely 'attributable' to the public service. Often, there are confounding factors (e.g. outcomes improve as society becomes more affluent), or a mix of two or more public services, or public and market services, play a part in achieving the outcome. For example, a hip operation that enables someone to walk provides a clear link between the output of the health service and an outcome, but good GCSE results may not be entirely attributable to schooling: support from parents, and possibly even private tuition, contribute. Development work to match outcome measures to spending on different services depends on assumptions, if they can be justified, that change in outcomes is attributable to public services. Techniques are needed to handle time lags, and to allow for imputation of future gains (where there is research evidence to justify this): the change in outcomes may be years, or decades, after the service has been delivered. Techniques are also needed to measure outcomes achieved by a combination of different public services: the 'Every Child Matters'

approach, discussed in section 5, will be one of the early areas for this. Measures should take account of negative impacts of public services on outcomes, as well as positive ones (a point made strongly in consultation seminars). This is an extremely difficult area, and perfect techniques (or perfect evidence on causality, in complex areas) will not be available. Nevertheless, the need to understand value for money from public spending is so important that further development work is a priority. ONS will develop further technical guidance, working with subject experts, aiming to find ways of using evidence on outcomes to supplement existing approaches to the measurement of public service output. This work will be linked to the separate ONS project on measurement of societal well-being. Results will be reported in productivity articles.

Methodological Issues in Quality Measurement

3.18. Three important methodological issues are discussed here in more depth:

- The relative importance of the different dimensions of quality
- Taking account of quality in output series by using 'value weights'
- The time lag between output and outcome.

Dimensions of Quality

3.19. A consumer may take separate account of different aspects of quality. They may put a very high weight on 'extensive hotel grounds', and a lower weight on 'range of choice at breakfast'. Similarly, measurement of public service output needs to consider whether all aspects of quality are equally important to users. This is compounded by the fact that (as with holidays), different aspects of quality matter to different people at different times. In principle, techniques like discrete choice experiments can be used to establish the relative importance of aspects of quality. The research to identify relative weights might be carried out in a sample representative of actual service users, but this may not be necessary or appropriate. The Personal Social Services Research Unit are working on this issue and are currently establishing the relative priorities of a sample of potential users of adult social care rather than actual users.

3.20. In the absence of research evidence about the relative importance of different dimensions of quality, some options are:

- Make a judgement about the most important dimension of quality (e.g. health gain, for NHS treatment) and use measures only in that area;
- Use indicators from more than one quality dimension and weight them as equally important;
- Use indicators from more than one quality dimension and seek expert opinion and carry out further analysis to determine how to weight them together. Sensitivity testing can be used to demonstrate the impact of alternative weights.

3.21. Any of these options can be justified, during development work, but the assumptions should be discussed explicitly, with recognition of the risks of error. In particular, it may be that a quality dimension is recognised, but no indicators are available. Another quality dimension may show one per cent improvement. Are there circumstances in which this might be used to judge that quality in the unknown

dimension has also improved by one per cent so overall quality improvement is one per cent? Or might it be more plausible to assume that the unmeasured dimension has shown no improvement and is equally important, so the overall quality improvement is 0.5 per cent? UKCeMGA considers that plausible assumptions about 'unknown' areas should be discussed explicitly, rather than relying only on indicators which can be measured but are known to give incomplete coverage of dimensions of quality. It is important to be transparent about how changes in quality are used to calculate the corresponding changes in output. This is an example of the use of 'professional statistical judgement', as proposed in *Establishing the Principles*; see paragraphs 3.36-3.40 below.

Taking Account of Quality in Output Series by Using Value Weights

- 3.22. Current output measures for the National Accounts weight the quantity of different public service outputs (e.g. primary school pupil attendances and secondary school pupil attendances) in proportion to their unit costs. One major strand of discussion in the consultation seminars was the merits of an alternative index, based on 'value weights' rather than 'cost weights'. This was recommended in the Atkinson Report and was also proposed as a major recommendation of the York/NIESR research on health output measurement. An index based entirely on value weights might show a different rate of growth from one based on cost weights. For example, high growth in low cost health treatments with major health benefits would contribute more to a value weighted index than to a cost weighted index.
- 3.23. Some respondents to the consultation felt that use of value weights would be the most important way of taking account of quality. Strictly, other methods for measuring quality change should also be used in conjunction with a value weighted output index – especially if it is not practicable to update value weights at all frequently. Many comments in the consultation supported the principle of value weights, though there were various counter arguments, such as comparability with other key economic statistics. Value weights are not used for public service National Accounts in any other country, and international comparability is important. A value weighted output series could be developed and published, even though it would not form part of the National Accounts.
- 3.24. Even those who agreed the principle recognised the difficulties in developing a complete set of value weights for any service. This would therefore be a long term development. There was opposition to the use of an index which mixes cost and value weights (see paragraphs 5.1-5.3). ONS will continue to work on the basis of cost weights. Where value weights are available, ONS will explore their use, particularly if this helps to test whether use of cost weights or use of value weights leads to markedly different measurements of the growth in output.

Time Lags

- 3.25. One important aspect of measuring quality that needs to be addressed is the issue of time lags. The outcome of a public service may not materialise until some time (maybe many years) after the provision of the service. For example, one of the purposes of education is for every child to be able to achieve economic well-being. Such a concept potentially lasts for the whole of the person's life, and is therefore impossible to measure first hand at the point at which the educational service is provided. A similar argument applies for the aim for children to be healthy – the true impact of this can only be measured over a lifetime.

- 3.26. The challenge is to develop realistic measures that approximate, in some sense, what we are trying to achieve. We can use existing data on lifetime outcomes (for instance a measure of economic well-being may use evidence on correlation between earnings of adults now in the workforce and their educational attainment). The question then is how to take into account the many other changes in society that have occurred over the years in order that we can apply evidence from today's adults to those children now who are in the education system. We expect their education to add to their future earnings, as individuals, but the effect will not necessarily be the same as for earlier cohorts.
- 3.27. As with other areas of development, UKCeMGA will consult widely with experts and practitioners in the development of methods.

Individual or Collective Services? What About Equity?

- 3.28. The conceptual model discussed so far is based on a service to an individual, where there is some implicit equivalent to a market price which takes account of different aspects of quality. Public services also have a 'collective' element. This is clear for services such as criminal justice, where a measurement framework is concerned with the benefit to society as a whole, not (just) the comfort of individuals who are arrested, tried or imprisoned. It is arguable that health also has a collective element: we all benefit from knowing an ambulance and intensive care bed are available for us if needed, even though we very seldom use them. We all benefit if the school system educates other people's children to be 'good citizens'.
- 3.29. Consultation seminars questioned why 'equity' was not part of the measurement of quality for health. Other performance frameworks sometimes include measures such as 'disparities in infant mortality between different social or ethnic groups.' If health is seen as an individual service, then the death of any baby is equally important, whatever the socio-economic status of the mother – this is why 'equity' has not been regarded as relevant in measurement of public services for the National Accounts. As a collective service, an objective of publicly funded healthcare is to reduce risks which are particularly high for some groups (e.g. risk of death for infants born to teenage mothers with low educational qualifications who smoke). Similar issues apply on inequalities within the education system.
- 3.30. Action to reduce inequalities may drive up unit costs, in a way which reduces measured productivity. If inequality is actually reduced, and this were to be regarded as a valid aspect of service quality, then this could be taken into account in the productivity calculation.
- 3.31. Further consideration is needed on whether reduction in inequality is relevant for measuring the quality of services which are (at least partly) 'collective'. This could be valid in productivity articles, even if it were not regarded as appropriate for the National Accounts.
- 3.32. There may be other ways in which the framework for measuring 'collective' services should differ from 'individual'. Where services have an 'insurance' element – valued because they are there when they are needed – measures of access may be important. Journey times to accident and emergency departments, for example,

matter to citizens who recognise they might need the service, not just to those who actually do use them.

Quality Measurement Framework (QMF) Project

- 3.33. UKCeMGA has secured additional funding from HM Treasury, and is working with partner organisations DH, DfES, the National Institute for Economic and Social Research, the Personal Social Services Research Unit at the University of Kent, and National Council for Voluntary Organisations, to develop a **quality measurement framework** for public services. This is a three year Invest to Save Budget project. The purpose is to develop effective but easily useable methodologies and procedures for measuring and assessing the value added of the relevant public services. They will constitute a tool so the service commissioning authorities can assess and monitor the performance of public, private sector or third sector providers. The tools need to link to the wider conceptual framework described in this strategy, but with practical application to front line services.
- 3.34. The QMF project will create entirely new mechanisms for more effective and efficient measurement and monitoring of third sector provision of public services, reducing the burden on the third sector whilst releasing cash through more efficient use of public funds to provide public services. The aim is to provide an effective tool that can be used by service commissioning authorities to make a reality of the policy intention. These general principles will be adapted on a case-by-case basis dependent on the individual service being investigated. The project aims to establish the differences in the quality and quantity of the provision of public services at the level of the individual provider. At the same time, use of the tool will reduce the administrative burden on service providers by concentrating on the essentials. The key requirement is that the single framework needs to embrace different users' needs, while still being uniformly applicable to measuring the output of all service providers, whether in the public, private or voluntary sectors. Success will allow more effective procurement and reduce the reporting burden on providers.
- 3.35. The project will research four areas of public service delivery in detail: care homes; knowledge and information services for adult social care; low-level social care interventions and preschool education. The initial report on the QMF project has been submitted to HM Treasury and will be available on the ONS website shortly. The project is due to finish in December 2009.

Methodological Quality

- 3.36. Development of measures of public service output that take proper account of quality is challenging. It is important to be realistic about the degree of accuracy which may be possible. The aim of the Atkinson Review, and the further developments discussed in the consultation *Establishing the Principles*, are to understand and reduce errors of measurement. Understanding comes first. Even if 'accurate' values are not knowable, it may be possible to say that current measures have consistent errors in a particular direction, or are volatile around a possible 'true' figure, or that error is getting smaller, or greater, over time. Understanding the reasons for and nature of error is the first step in measuring error. The Atkinson Review pointed out that failure to measure quality of public services, or possible change in quality over time, is an important source of error.

- 3.37. *Establishing the Principles* set out an approach for using ‘professional statistical judgement’ to evaluate possible methods of output measurement, based on a clear framework for quality measurement. Consultation endorsed the need for professional judgment, but there were concerns that rigour might be compromised by data limitations or by the agenda of those closest to service delivery. It was agreed that the standards of ‘good enough’ depended on the purpose of measurement and the use to which it would be put, with National Accounts requiring a particularly stringent approach, given the requirements for international comparability. The need for transparency was emphasised.
- 3.38. UKCeMGA will therefore continue to develop and document methods carefully, engaging stakeholders - experts, practitioners and the international community. In addition to development work led by UKCeMGA, other developments may be taken forward by government departments, academics and other interested bodies. We propose that developmental work carried out by ourselves or others will be available for mutual scrutiny and peer review as part of the development process. UKCeMGA will aim to set out clear standards for public sector output measures. Then we can categorise methods as being suitable for use, in development, needing more work etc. One possible set of standards is as follows:
- **Concepts** - the extent to which the particular concept is accepted, and the closeness with which it is being measured
 - **Techniques** - the relevance and robustness of methods employed
 - **Data** - issues such as relevance, representativeness, consistency, timeliness and completeness.

Different developments will have different balances of those three aspects. So for instance, a useful technique might be developed, but without having appropriate data to which to apply it. The interaction of the factors also needs to be considered: in some cases, poorer quality data may have little impact on final estimates, so their use would be more acceptable than if the impact were larger. It is therefore important to develop methods in ways which allows the effects of particular issues to be tested.

- 3.39. UKCeMGA will aim to win general acceptance for methodological standards that can be used in this way, to help the pace and coherence of development and interpretation.
- 3.40. Finally, methodological developments will be based in the principles in the National Statistics Code of Practice and its associated protocols. Relevant guidance is:
- ‘National Statistics will gain public trust through being produced using objective and transparent methods’ and ‘...will be fit for purpose and of high quality.’
 - ‘The selection of statistical sources, methods and procedures is a professional responsibility and will be based on scientific principle and best international practice, taking into account cost implications to government and providers’.
 - ‘Knowledge of National Statistics will be sustained by quality measures, regularly published.’

Conclusions for Methodology:

What is now established:

- A quality measurement framework is required to develop authoritative, credible and transparent estimates of public service output.
- Key concepts for the framework include defining the right unit of output, identifying the appropriate dimensions of quality, applying the right analytical techniques and considering the capacity for benefit of individual users.
- Three particular methodological issues need attention: the relative importance of individual quality dimensions, using value weights instead of cost weights, and consideration of the time lags between outputs and outcomes.
- Professional statistical judgement is needed throughout the development of the quality measurement framework.

Priorities for immediate work:

- Development of the quality measurement framework that is backed by clear analysis and documentation of the methodological issues involved.
- Establishing a set of robust methodological principles to support the quality measurement framework.

Gaps and challenges:

- Establishing systematic sets of value weights to measure public service output (for potential use in productivity articles and satellite accounts).

4. The Measurement of Public Service Output: Accounting for Private and Public Sector Complementarity

- 4.1. Complementarity is a complex issue. Principle C from the Atkinson Review proposed that 'Account should be taken of the complementarity between public and private output, allowing for the increased real value of public services in an economy with rising real GDP'. One example is that individuals value health and education services more today than twenty years ago, since good health allows individuals to enjoy more valuable leisure activities. More generally, education supports economic growth, and healthcare reduces time off work due to sickness. The private sector could also benefit from public sector output: a delivery firm may become more efficient when the government provides improved infrastructure or regulation that benefits their business.
- 4.2. The issue for consultation was whether this sort of linkage should be taken into account by an explicit adjustment to the measure of public service output. Presentation of the arguments in favour of including an adjustment to public service output to reflect its contribution to growth in the market sector, in a way which can be followed by a lay audience, failed to convince most people at consultation, though appealed intuitively to others. Most of the more technical responses opposed the argument, as published in the response paper. For these reasons, UKCeMGA decided to complete the consultation by asking a small group of expert economists, most of whom specialise in public sector economics, to review all the issues and prepare a report.
- 4.3. An independent economic advisory panel was set up, chaired by Robert Chote, who is Director of the Institute for Fiscal Studies and a member of the UKCeMGA Advisory Board. The panel were invited to submit written responses to a structured questionnaire about whether measures of public service output should be adjusted to take account of complementarity. Questions separated three contexts: the National Accounts, productivity articles, or other types of analysis. One key issue was whether any rise in the value of public services during times of economic growth should just be regarded as a relative price effect, or whether there were unmeasured effects on the volume of public services.
- 4.4. The panel met once, Robert Chote then prepared a report which was agreed by all members. The reports, and names of members, are at Annex D. The panel recommended:
1. Where there is ambiguity as to the distinction between price and output effects, the presumption should be that they are relative price effects rather than changes in volume. Adjustments based on public/private sector complementarity should only be made where there is a compelling argument in their favour.
 2. Measures of public service output incorporating the 1.5 per cent earnings adjustment should be dropped from UKCeMGA productivity articles.
 3. UKCeMGA should consider adjusting the output of the protective services to reflect changes in the volume of assets protected.

4. ONS should develop an analytical framework for welfare measures broader than GDP, for which adjustments for private/public sector complementarity specific to particular services may be appropriate.
5. Welfare measures for particular services should be published in the productivity articles, once the analytical framework is fully developed and explained.

UKCeMGA will work on the basis of these recommendations.

5. Applying the Quality Measurement Framework to Education Output

- 5.1. In October 2005, ONS published the first education productivity article which presented figures from four options for quality adjustment based on exam attainment, using GCSE and Key Stage Progress results. The methods were set out in more detail in a paper published by DfES earlier in October 2005. Key issues were set out in the education consultation paper published by ONS in November 2006.
- 5.2. This strategy is mainly concerned with methods of quality adjustment. The consultation agreed that methods should take account of the different education systems in the UK rather than just focus on the English system. Consultation also highlighted ideas for extending the basic quantity measure of schools output, which is based on pupil attendance years. It was thought desirable to add to this by counting extra curriculum activities such as sport, music and drama, and to take proper account of time spent by children on school premises outside core teaching hours, as the 'extended schools' policy is implemented.
- 5.3. The coverage of the current output measure was also questioned. In line with National Accounts definitions, public service education includes sixth forms at schools but not the equivalent education provided by sixth form colleges and further education colleges. A more inclusive approach would fit better with education policies for 14 – 19 year olds. Quality measures, including examination attainment, should be used consistently across the different types of establishment. UKCeMGA will undertake development work to widen its approach to cover a wider range of educational output, rather than being limited to the precise National Accounts definitions. Progress will be shown in future productivity articles.
- 5.4. The current use of cost weights in the education output measure was accepted during consultation without discussion. A case might in principle be made for using value weights, e.g. reflecting the importance of early years' education for subsequent progress: this might be explored in productivity articles, though not for the National Accounts.
- 5.5. It was noted in consultation that work so far had considered only one dimension of quality of school output, i.e. pupil attainment measured through examinations. Other aspects of quality were recognised: pupils' sense of enjoyment and wellbeing; schools as a healthy environment able to help children avoid obesity; success of schools in developing future citizens who avoid problems such as crime and drugs. A wider approach to measuring quality of school output is now proposed, linked to the outcome framework published in the 2003 Green Paper *Every Child Matters*. This approach enables ONS, DfES and the Devolved Administrations to explore a range of indicators being developed for other purposes, but ONS would need to be sure that its use of quality measures fits a wider framework which is accepted by impartial experts, not open to criticism as being short term or politically driven.

Quality Measures Based on Examination Attainment

- 5.6. An ideal method based on examination attainment would measure the value added from each year of education to each cohort of children, for their schooling up to the age of 16, or older. The Foundation Stage Profile currently collected by schools in

England was introduced on the basis that it was not suitable for value-added calculations. It is therefore not possible to produce a value added measure, which would assess directly educational progress from entry to primary schools. The Curriculum Evaluation and Management Centre at the University of Durham do collect baseline information on school entry from a number of schools. There will be further assessment of whether a national quality adjustment method could be based on this sample. It may also be possible to develop a method using the time series for Key Stage results, looking particularly for methods which can be linked to pupil and spending information used as the components of the volume measure. For example, a quality adjustment for primary schools, using information on progress between Key Stage 1 (age 7) and Key Stage 2 (age 11) might be linked to the pupil numbers and unit costs for primary schools

- 5.7. The consultation agreed that an adequate short term measure could continue to be based on GCSE results. A method using the 'average point score' (APS) was preferred to one based on the proportion of pupils with five or more good GCSE passes because it better represents the experience of all pupils. An average point score for GCSE exams is constructed using points for different grades. This does not make full use of the actual marks scored by candidates but reflects an assumed weighting for different marks in different ranges. The weightings were changed in 2004. The GCSE measure also has 'floor and ceiling' effects: it gives no indication of attainment for the small proportion of 16 year olds who do not pass any GCSEs, and does not give extra points where pupils take more than eight GCSEs, so does not fully measure change in the educational prowess of those with highest attainments. Also, it is a lagged indicator which measures the experience a single cohort has acquired over 11 years, and cannot reflect any emerging changes in the education experience of younger cohorts.
- 5.8. The quality adjustment for the education method adopted for the National Accounts in 1998 was based on the APS score, using a particular algebraic model, which is only one of a number of possibilities. ONS is working with DfES to establish which method is most plausible, and will publish its conclusions in the forthcoming Education Productivity article. It is also important to use exam results from each part of the UK, particularly Scotland which has a distinct examination system.
- 5.9. In the medium term, ONS will work with DfES and other education departments to attempt further improvements in the GCSE based measure, also taking into account the different exam systems within the UK. Consideration will be given to replacing the current points scores, for different GCSE grades and equivalents, with weights for different levels of attainment which are based on future earnings, following a method developed by the Scottish Executive. The consultation suggested that Maths, English and perhaps IT are key GCSEs and more weight should be given to these. Other work will develop approaches for measuring attainment after 16, including evidence from sixth form colleges and colleges of further education.

Other Quality Domains

- 5.10. Education is about more than just passing exams. Education helps young people obtain the skills they need to participate fully in society. The 'Every Child Matters' (ECM) framework captures these skills and the wider outcomes of education under five outcomes:
 - be healthy;

- stay safe;
- enjoy and achieve;
- make a positive contribution; and
- achieve economic well-being.

- 5.11. Although these outcomes cannot be achieved solely by education, education has a central role to play: for example, by encouraging healthy eating in schools, offering care and emotional support, giving opportunities to participate in sports, art, music, drama, and developing social and communication skills. Children and young people learn and thrive when they are healthy, safe and engaged; and the evidence shows that educational achievement is one of the routes out of poverty. Each of the outcomes is also affected by many factors other than schooling – families, but also other public services, including children’s social services, healthcare and youth justice, and many aspects of the market economy. So in measuring outcomes for children, it is important to consider carefully what should be attributed to schools, rather than other factors.
- 5.12. ONS, with DfES and the Devolved Administrations, will develop a quality measurement framework based on the ECM outcomes which can be used to adjust the volume measure of education output. Key issues for development work include: whether all five outcomes have equal weight; which indicators map to which outcomes (e.g. pupil attainment could be seen as a proxy to both “enjoy and achieve” and ‘achieve economic well-being’); and how far change in any outcome indicator is attributable to schools.
- 5.13. The consultation agreed that pupil attainment is a key aspect of quality, but that other dimensions of quality were also important. In continuing to use the available data on GCSE attainment, in advance of a fuller framework for other outcomes, consideration is needed on how to allow for known unknowns – as set out in paragraphs 3.19-3.21 above. The judgement should be taken in the context of other acknowledged sources of error in using the education attainment measure, such as the floor and ceiling effects. ONS will set out its conclusions, with reasons, in the forthcoming education productivity article.

Other Development Work

- 5.14. ‘Capacity for benefit’ is important in comparing educational attainment, or other outcomes, across a cross section of pupils or of schools. It is less clear whether there are relevant changes in ‘capacity for benefit’ from education between successive cohorts. Factors such as inclusion of more children with special needs in main stream schooling, or a change in the proportion of children for whom English is not a first language, could be considered. Research on wider social, behavioural and health issues within schools is important to understand the context for school output. One aspect of the quality of the school system is the extent to which appropriate provision is made for every individual child; and quality measures for access and adequacy of support services might also be useful.
- 5.15. Consultation did not support development of quality adjustments based on measures of class size. Evidence from Ofsted inspection reports was recognised as valuable, but not suitable for measuring change over time since the inspection system and grading have changed several times. There was interest in the use of

the Programme for International Student Assessment (PISA) results, particularly for comparisons between countries, but it was recognised that the focus of PISA changes every time, and tests are not carried out every year, so this is not a useful base for a time series. The consultation also commented that PISA has its own costs, in terms of pupil and teacher time as well as financially, so taking a larger sample or testing more frequently were not regarded as easy options.

5.16. Measuring quality of under five education is being taken forward through the QMF project (paragraphs 3.33-3.35).

Conclusions for Education:

What is now established:

- The quantity output measure should be extended to include extra curriculum activities as well as lessons, which are currently proxied by pupil attendance days, and to include extended school days.
- The quantity output measure should be extended to include attendances (at least up to age 18/19) at sixth form colleges and FE colleges.
- More development work is needed to consider whether Key Stage results can be used for a measure of the quality of education at primary school level.
- In the short term, GCSE results should be used for a quality adjustment based on the average points score not the threshold of 5 good passes.
- The pupil attainment quality measure should be developed to incorporate results from A levels and diplomas up to age 19, at schools and sixth form colleges; there should be exploration of alternative weights based on the links from exam results to potential average earnings, or higher weight for qualifications in Maths, English and IT.
- There is a need to extend the number of quality domains more widely than exam attainment.

Priorities for immediate work:

- Reviewing the best model for a quality adjustment based on GCSE results.
- Establishing the availability of data to extend the quantity measure.
- Establishing a quality measurement framework based on the ECM outcomes and identifying suitable quality indicators for the outcomes.
- Further research on alternative models for the pupil attainment quality measure.

Gaps and challenges:

- Fundamental problem of measuring education quality only at age 16, in absence of baseline data to measure value added in earlier years.
- Judging how much of the change in any indicator is attributable to schools, rather than other factors.

How to determine the weights of the different outcomes in a single quality measure.

6. Applying the Quality Measurement Framework to Healthcare Output

- 6.1 In February 2006, UKCeMGA published the second health productivity article which set out some possible approaches to including quality adjustments in the health output measure, based on publications in December 2005 by the Department of Health (*Accounting for Quality Change*) and by a consortium of researchers from the University of York Centre for Health Economics and the National Institute for Economic and Social Research (York/NIESR). The York/NIESR research was commissioned by DH and was developed in parallel with the Atkinson Review; the DH paper summarised the York/NIESR research and included a number of additional possible approaches to quality adjustment. These proposals were summarised in the health consultation paper issued by UKCeMGA in December 2006.

Value Weights for Healthcare Output

- 6.2 There is a well-established healthcare output index, based on unit costs and counts of a large number of separate NHS activities. The York/NIESR research strongly recommended that cost weights should be replaced with value weights, based mainly on the value of health gain given by each type of treatment. Other aspects of quality could be included in the value weight, if sufficient information were available. In practice, the researchers could identify value weights for only a limited set of hospital treatments. DH demonstrated a calculation of a value weight for statin prescriptions – a widely used drug which has a significant effect in reducing the chance of heart attack or stroke.
- 6.3 There was strong support, particularly from health economists, for the principles of value weights. However, most also believed that it would be very difficult to estimate such weights across the full range of healthcare in order to construct a value weighted index. There were strong doubts about the validity of a mixed ‘value and cost’ weighted index, where individual value weights were substituted into a cost weighted index as they were developed. There was particular concern that development work might be biased towards treatments where value was higher than cost, rather than the reverse; participants hoped that the National Institute for Health and Clinical Excellence (NICE) would be able to help an impartial development programme which included treatments where cost exceeded value. There was also concern that value weights introduced an element of consumer surplus into the index (if patients were getting a bargain from statins, the higher price that patients would have willingly paid was not relevant for National Accounts, which should be based on the equivalent of the market price); and also concern that productivity gains developed by the pharmaceutical industry were being included inappropriately into measures of NHS productivity.
- 6.4 UKCeMGA will continue to base its healthcare output measures, for the National Accounts and for productivity articles, on an index which uses cost weights alone. UKCeMGA will discuss with health administrations and NICE the scope for further exemplification of the impact of using value weights, aiming for a balanced approach. Such work could be published in productivity articles (or by other bodies) to continue the interesting debate on concepts and interpretation. It is clear that a claim that an index with mixed value and cost weights ‘reduces error’, compared with a cost weighted index, would be rejected by many observers. Nevertheless, it is

difficult to see how a complete value weighted output index could be developed without regular publication of interim steps, and discussion of the basis of calculation. UKCeMGA will support such developments, because of their conceptual importance, but this will not be a major priority within the current work programme.

Quality Dimensions, Indicators and Methods

- 6.5 The four main dimensions or domains of healthcare quality proposed by the Atkinson Review were broadly accepted by most people at consultation. One of the four domains, speed of access, was considered to have two distinct parts, each of which already formed part of other domains. One part is the extent to which quicker access achieves better results: this is part of saving lives and extending life span. The other part was convenience to the patient: this is part of quality of patient experience. UKCeMGA therefore considers the main domains of healthcare quality to be:
- saving lives and extending life span;
 - preventing illness and mitigating its impact on the quality of life;
 - quality of patient experience.
- 6.6 ONS is aware of other development work by DH on a quality framework and will discuss ways of mapping from the three domains above to other systems, if necessary. There was also interest in equity (of health outcome and of access) as a quality domain: see paragraphs 3.26-3.32 above. There was strong emphasis on the importance of measuring the effectiveness of the NHS in 'preventing illness', but recognition of difficulties in measuring illness which did not happen, and in assigning outcomes that may not be realised until some time after the initial treatment. Participants noted that some activity in primary care amounts to reassurance, rather than health improvement, in cases where patients actually have no health problems.
- 6.7 Most commentators felt that dimensions of quality based on health gain – saving lives, mitigating effects of illness – were more important than patient experience. There were doubts about the approach in the DH publication, which gave equal weight to 'health gain' and to 'patient experience'. ONS agree that patient experience should have a lower weighting than health gain, though at present there is no research base to say precisely what the relative weights should be. Further research, and/or expert consultation, is needed to justify a particular basis for relative weights, but it would be valid to make some simple assumptions, transparently, rather than ignore the 'patient experience' dimension of quality altogether – or to assume it has equal importance with health gain.
- 6.8 Consultation welcomed the work by York/NIESR and DH to identify specific quality indicators related to these domains, such as survival rate for 30 days after hospital admission, though this did not cover other important aspects of healthcare quality. Analysis of health gain from primary care, in terms of blood pressure control, was an interesting start. There were some doubts about work done so far on quality indicators for waiting times and patient experience.
- 6.9 For all the existing proposals, consultation provided many suggestions for further development, either in terms of improving the method or acquiring better data. ONS will continue to ensure that its estimates of output and productivity are accompanied

by careful explanations of known and measurable error and discussion of other types of error.

- 6.10 There were concerns about difficulties in attributing outcomes (including health status) to healthcare spending. Participants thought it very important to understand the impact on health status from NHS activity and build this into output and productivity estimates. One way forward may be to focus on mortality from conditions amenable to health care, rather than the 'all deaths within 30 days of hospital admission' data used by York/NIESR, since death in hospital is (for some patients) a natural end of life. For the longer term, UKCeMGA will discuss with health administrations and information centres how best to begin building up an evidence base on attribution of health outcome from treatment for different conditions.
- 6.11 Consultation expressed a desire for better basic methods of measuring healthcare activities which linked separate episodes of treatment for the same patient. Data sets to measure patient health outcomes after treatment, routinely, would be very valuable.
- 6.12 Study of changes in health status after different treatments also needs to take account of case mix. For example, if waiting times reduce, patients may have less severe symptoms when they are treated, and so receive less benefit from treatment. Such differences over time in morbidity should be taken into account, though there were doubts whether routine NHS data would allow this. (This 'case mix adjustment' factor is similar to the 'capacity for benefit' point discussed earlier).

Conclusions for Healthcare:

What is now established:

- Productivity articles should continue to use results from recommendations made by York/NIESR (for survival, health gain and waiting times), but with consideration of an 'avoidable mortality' adjustment rather than data for all deaths within 30 days of hospital admission.
- Productivity articles should use developing work from DH using new data (QRESEARCH) for outcomes in general practice.
- Patient experience measures should be used for quality adjustment, subject to further technical assessment, but with a weighting which is lower than for health gain.
- A mixed 'cost and value weight' approach will not be used in healthcare productivity analysis, though developments in identifying value weights will be reported.

Priorities for immediate work:

- Updating and refining previous calculations of healthcare quality change, working with DH, for publication later in 2007.

Gaps and challenges:

- Healthcare activity measures are a limited proxy for a 'unit of healthcare output' which links different activities involved in treating the same patient.
- Techniques to take fuller account of health prevention within the output measure, including measures of change in impact on outcomes.
- Limited range of evidence on impact of healthcare on its intended outcomes.

7 Applying the Quality Measurement Framework to Other Public Services

7.1 The consultation discussed education and healthcare in more detail, but the broader framework for quality measurement applies also to other public services. The other areas where ONS has been working so far are:

- Adult social care
- Children's social services
- Administration of social security and benefits
- Criminal justice
- Fire services
- Civil justice
- Defence

7.2 The range of services could be further extended to areas such as immigration services, tax administration, social housing, and libraries. Work on defence is only just starting, and the priority is to make progress on all the services above.

Adult Social Care

7.3 Development of a quality framework for adult social care is fairly advanced and the research for this framework forms the cornerstone of the QMF project (paragraphs 3.33-3.35). As such it fits comfortably within the general strategy outlined in this paper. Capacity for benefit is a particularly important consideration for adult social care: the same level of service will not necessarily achieve the same benefit for each individual, depending on their needs. Change over time in the dependency of users is significant, and is a reason for rising unit costs. Use of objective measures of this factor, as well as quality of service, is important in understanding productivity change.

Children's Social Services

7.4 Output measurement and quality adjustment for children's social services is less advanced but a report by the Centre for Child and Family Research at Loughborough University investigates ways of improving the measurement of the output of Children's Social Services in the National Accounts. In line with the recommendations from the Atkinson Review, the output measure has been a cost-weighted index, comprising four activity indicators for services provided to looked after children:

- number of child-days in children's homes
- number of child-days in secure accommodation
- number of child-days in foster placements
- number of child-days in other placements

7.5 The output of the remainder of Children's Social Services' activities is measured indirectly by the total expenditure on these activities. The research is looking into how output might be further improved by:

- broadening the range of activities for which output is measured directly rather than via expenditure;
- broadening the methodology by which these indicators can be weighted together to form an index; and
- adjusting output for any changes in the quality of service provision over time.

7.6 Children's social services outcomes form a large part of the Every Child Matters agenda and there will be overlaps with the education agenda. Where possible these two agendas will be taken forward together under the Every Child Matters outcome framework.

Administration of Social Security and Benefits

7.7 In social security administration, the measurement of individual aspects of quality is well advanced (based on accuracy and timeliness of payments) but work now needs to be undertaken to bring these together into a coherent framework which covers all applicable dimensions of quality.

Criminal Justice

7.8 Work on measuring the output and productivity of the criminal justice system (CJS) is relatively new and is not as far developed as other areas of output measurement. Work in this area will be taken forward in two ways. The first strand is based on the Administration of Justice Framework (developed by the Home Office, Office for Criminal Justice Reform, Department for Constitutional Affairs and the Crown Prosecution Service) to calculate the output, and consequently productivity, of the CJS as three separate components:

- i. detection through to sentencing (police, Crown Prosecution Service and criminal courts);
- ii. administering the sentences imposed by the courts;
- iii. non-investigative work (this would include activities related to crime prevention, for example, much of the community support policing work and police patrols).

7.9 The second strand is longer-term work that aims to develop an econometric approach for estimating the effects of these activities on crime rates. For both strands of work, the underlying principles of the quality measurement framework will be applied, with particular development of methods suitable for 'collective' rather than 'individual' services. Use of qualitative measures of people's confidence and satisfaction in the criminal justice system could be an important source, for example using the Police Performance Assessment Framework (PPAF) time series on victim satisfaction. (As with health care, people would generally much prefer to avoid *any* contact with the criminal justice system – but it is still possible to measure the way they regard various aspects of the service when they do come into contact.)

Fire Services

7.10 A similar approach will be applied to the fire services, where again progress in measuring the output and productivity of this service is less developed. The ideal measure of total output would be the value of damage mitigated by fire and rescue

services. It has been recommended that we look at the value of the fire and rescue activities, whether responding to fires, road traffic accidents or raising awareness of fire safety when looking in to the damage that these services might mitigate. Quality in the fire and rescue services is essentially a measure of the effectiveness of the service provided, in terms of the final outcomes. In terms of fire safety activities, a measure of quality should reflect how good each activity was at raising awareness of fire safety, changing people's behaviour and, ultimately, reducing the number of fire starts.

Civil Justice

- 7.11 Direct measures of output for county courts are already used in ONS, although these measures have significant limitations. A paper by the DCA 'Measuring output and productivity in the civil and family court system' sets out the methodology for improving the direct output measure for civil and family courts, probate registries and a part of the Chancery division (high court). The methodology proposed uses quantity data weighted by unit costs of activities (calculated for the previous year) to estimate value of real output. The output measure proposed has limitations in that it does not cover some important Ministry of Justice output (e.g. alternative dispute resolution), nor does it reflect changes in quality of service provided by the civil and family court system. These are recognised as issues that need urgent consideration in the next stage of development.

Defence

- 7.12 Work is only just commencing on measuring defence output and productivity, moving away from the existing output=inputs convention.

8 Next Steps

8.1 ONS will work to implement this strategy by:

- Working closely with relevant government departments, including the devolved administrations, to ensure relevant expertise is applied within a coherent framework;
- Discussing the quality measurement framework, and its use, with other bodies with roles in public service accountability and regulation, to share expertise;
- Continuing to improve productivity methods not discussed in this paper, including accurate identification of spending for each programme, use of best techniques to deflate for pay and price change, and widening the range of activities and unit costs to construct the output index in each area.
- Making use of research, commissioned by ONS, by departments or published independently;
- Publishing productivity articles; current plans are for these four articles to be published in early autumn 2007:
 - Education services
 - Health services
 - Administration of social security
 - Adult social services.

Annexes

A References

Atkinson A B (2005), *Atkinson Review of measurement of government output and productivity for the National Accounts: Interim Report*, July 2004, available at <http://www.statistics.gov.uk/about/data/methodology/specific/PublicSector/atkinson/default.asp>

Department for Education and Skills (2005), *Measuring output from the Education System*, available at <http://www.dfes.gov.uk/research/programmeofresearch/index.cfm?type=5&keywordlist1=0&keywordlist2=0&keywordlist3=0&andor=or&keyword=education+output>

Department of Health (2005), *Health output and productivity: Accounting for quality change*, December 2005, available at http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4124266

Department for Work and Pensions (2005), *An analysis of DWP Productivity 1997/98 – 2007/08*, available at <http://www.dwp.gov.uk/asd/asd5/rrs2006.asp>

Netten, Ann et al (2006), *Measuring Personal Social Services Outputs for National Accounts: Services for Older People*, Final Report, PSSRU Discussion Paper 2267/3, January available at <http://www.pssru.ac.uk/abstracts.php?id=DP2267/3>

ONS 2004, *Public Service Productivity: Health*, Office for National Statistics, October 2004, available at <http://www.statistics.gov.uk/CCI/article.asp?ID=1030&Pos=2&ColRank=1&Rank=224>

ONS 2005, *Public Service Productivity: Education*, Office for National Statistics, October 2005, available at <http://www.statistics.gov.uk/cci/article.asp?id=1290>

ONS 2006a, *Public Service Productivity: Health*, Office for National Statistics, February 2006, available at <http://www.statistics.gov.uk/CCI/article.asp?ID=1030&Pos=2&ColRank=1&Rank=224>

ONS 2006b, *United Kingdom National Accounts The Blue Book 2006*, Office for National Statistics, August 2006, available at http://www.statistics.gov.uk/downloads/theme_economy/BlueBook2006.pdf

ONS 2006c, *Sources and methods for Public Service Productivity: Health article*, Office for National Statistics, February 2006, available at <http://www.statistics.gov.uk/cci/article.asp?id=989>

ONS 2006d, *Establishing the Principles Consultation*, Office for National Statistics, September 2006, available at <http://www.statistics.gov.uk/about/data/methodology/specific/PublicSector/output/principles.asp>

ONS 2006e, *Education Consultation*, Office for National Statistics, November 2006, available at http://www.statistics.gov.uk/about/data/methodology/specific/PublicSector/output/downloads/UKCeMGA_Education.pdf

ONS 2006f, *Health Consultation*, Office for National Statistics, December 2006, available at http://www.statistics.gov.uk/about/data/methodology/specific/PublicSector/output/Health_consultation.asp

York/NIESR 2005, *Developing new approaches to measuring NHS output and productivity*, York University and National Institute of Economic and Social Research, available at www.york.ac.uk/inst/che/pdf/tp31.pdf

Useful Links

Written responses to the three consultations and the related seminar summaries are available via the following links:

- Establishing the Principles Consultation:
<http://www.statistics.gov.uk/about/data/methodology/specific/PublicSector/output/principles.asp>
- Education Consultation:
http://www.statistics.gov.uk/about/data/methodology/specific/PublicSector/output/Education_consultation.asp
- Health Consultation:
http://www.statistics.gov.uk/about/data/methodology/specific/PublicSector/output/Health_consultation.asp

B Written Consultation Code of Practice

B.1 The Cabinet Office has published a *Code of Practice for Written Consultations* to guide departments' activities in this area. The full code is available at www.cabinetoffice.gov.uk/regulation/consultation/consultation_guidance/index.full.asp. The Code of Practice includes six key criteria to be applied to all departmental consultations, set out below.

Consultation Criteria:

- Consult widely throughout the process, allowing a minimum of 12 weeks for written consultation at least once during the development of the policy.
- Be clear about what the proposals are, who may be affected, what questions are being asked, and the timescale for responses.
- Ensure the consultation document is clear, concise and widely accessible.
- Give feedback regarding the responses received and how consultation process influenced the policy.
- Monitor the department's effectiveness at consultation, including through the use of a designated consultation co-ordinator.
- Ensure your consultation follows better regulation best practice, including carrying out a Regulatory Impact Assessment if appropriate.

B.2 If you feel that this consultation does not meet these criteria, please contact:

Peter C Smith
Chairman, UKCeMGA Advisory Board
University of York Centre of Health Economics
University of York
Heslington
York
YO10 5DD

Confidentiality Disclosures

B.3 Information provided in response to this consultation, including personal information, may be published or disclosed in accordance with the access of information regimes (these are primarily the Freedom of Information Act 2000 (FOIA), the Data Protection Act 1998 (DPA) and the Environmental Information Regulations 2004). If you want information that you provide to be treated as confidential, please be aware that, under FOIA, there is a statutory Code of Practice with which public authorities must comply and which deals, amongst other things, with obligations of confidence. In view of this, it would be helpful if you could explain to us why you regard the information you have provided as confidential. If we receive a request for disclosure of the information, we will take full account of your explanation, but we cannot give an assurance that confidentiality will be maintained in all circumstances.

B.4 An automatic confidentiality disclaimer generated by your IT system will not, of itself, be regarded as binding on the Department. The Department will process your personal data in accordance with the DPA, and in the majority of circumstances, this will mean that your personal data will not be disclosed to third parties.

Freedom of Information Contact

Correspondence and Enquiry Unit
Freedom of Information Section
HM Treasury1 Horse Guards Road
London SW1A 2HQ

Fax: (+44) (0)20 7270 4681

E-mail: public.enquiries@hm-treasury.gov.uk

C Principles from the Atkinson Review Final Report

The following main principles cover the direct measurement of output, the measurement of inputs, and the measurement of productivity.

- **Principle A:** the measurement of government non-market output should, as far as possible, follow a procedure parallel to that adopted in the National Accounts for market output.
- **Principle B:** the output of the government sector should in principle be measured in a way that is adjusted for quality, taking account of the attributable incremental contribution of the service to the outcome.
- **Principle C:** account should be taken of the complementarity between public and private output, allowing for the increased real value of public services in an economy with rising real GDP.
- **Principle D:** formal criteria should be set in place for the extension of direct output measurement to new functions of government. Specifically, the conditions for introducing a new directly measured output indicator should be that (i) it covers adequately the full range of services for that functional area, (ii) it makes appropriate allowance for quality change, (iii) the effects of its introduction have been tested service by service, (iv) the context in which it will be published has been fully assessed, in particular the implied productivity estimate, and (v) there should be provision for regular statistical review.
- **Principle E:** measures should cover the whole of the United Kingdom; where systems for public service delivery and/or data collection differ across the different countries of the United Kingdom, it is necessary to reflect this variation in the choice of indicators.
- **Principle F:** the measurement of inputs should be as comprehensive as possible, and in particular should include capital services; labour inputs should be compiled using both direct and indirect methods, compared and reconciled.
- **Principle G:** criteria should be established for the quality of pay and price deflators to be applied to the input spending series; they should be sufficiently disaggregated to take account of changes in the mix of inputs; and should reflect full and actual costs.
- **Principle H:** independent corroborative evidence should be sought on government productivity, as part of a process of 'triangulation', recognising the limitations in reducing productivity to a single number.
- **Principle I:** explicit reference should be made to the margins of error surrounding National Accounts estimates.'

D Chairman's Final Report for the Economics Advisory Panel

Measuring the Output and Quality of Our Public Services

Chairman's report of an Economics Advisory Panel (EAP) meeting on the appropriateness of accounting for public and private sector complementarity when measuring the output of public services

EAP members: Tony Atkinson
Robert Chote (Chair)
Graham Loomes
Peter Smith
Martin Weale

UKCeMGA attendees: Fathimath Ali
Stephen Almond
Joe Grice
Aileen Simkins

Background

1. In its recommendations for the measurement of the output of public services in the national accounts, the Atkinson Review argued in 2005 that: "To a significant degree the output of government services rises with the real value of private assets and incomes... Unless it is taken into account there is serious risk that the output of the public sector will be understated. This may be summarised in Principle C: account should be taken of the complementarity between public and private output, allowing for the increased real value of public services in an economy with rising real GDP".
2. In practice, UKCeMGA has applied this "complementarity principle" by publishing alternative series for the output of key public services, some of which increase the volume of output by an extra 1.5 per cent a year to reflect average earnings growth. In the case of health, for example, this is justified on the grounds that "health becomes increasingly valuable in a growing and increasingly productive economy". But UKCeMGA has warned in its publications that "this adjustment should be used cautiously pending further debate".
3. In press and other comment, the earnings adjustment has been more controversial than the other recommendations made by the Atkinson Review, to capture quality improvements more effectively. This is in part because it is the most significant quantitatively.
4. In the light of external reaction, UKCeMGA and its advisory board have been concerned that credibility is being undermined by the presentation of the earnings adjustment as an option, without a strong indication of whether it enjoys official backing. If UKCeMGA is confident that the adjustment should be used in national accounts measures of public service output, then it should be explained and defended more robustly than has been the case to date. If not, then it would be better to remove this measure from the productivity articles and other publications unless the weight of argument in its favour becomes such that it can be explained and defended properly.

5. The recent consultation process undertaken by UKCeMGA has revealed widespread doubts about the merits of the earnings adjustment, but without sufficient agreement on how to proceed to justify its removal or replacement in the productivity articles.
6. With UKCeMGA due to publish a Strategy Paper on the methodology for public service output measurement in July 2007 - and a fresh round of productivity articles to follow – it was felt important to reach closure on this issue. The EAP was therefore convened to make firm recommendations on the use of the earnings adjustment in productivity articles and on other avenues to address the complementarity of public and private sector output.

Complementarity and the National Accounts

7. The panel agreed that the central problem in measuring the contribution of public services in the national accounts is to distinguish between changes in volume and changes in price. This is difficult enough with private services, where contracts are not specified completely, but arises particularly acutely in the case of public service outputs where there is no market transaction.
8. The balance of opinion is that, where there is ambiguity as to the source of the change, it should be treated as a relative price effect rather than as an increase in volume. In view of this, the panel recommends that an adjustment based on public/private sector complementarity should only be made where there is a compelling argument.
9. In the case of education, the panel did not feel that an adjustment should be made for real earnings growth. It recognised that this is likely to mean that the volume of education output, as measured, would not keep up with the volume of inputs in the long-run.
10. The panel agreed that in some cases there is a compelling case to take account of growth in the private sector, notably in the protective services, where the public sector output may be seen as increasing with the volume of private sector goods protected. Such adjustments should be introduced on a case by case analysis, based on rigorous research and full consultation.
11. The panel noted that discussion of this issue underlined the importance of considering carefully whether some productivity growth currently attributed to the private sector in the national accounts should be attributed to the public sector – and vice versa. But the panel recognised the difficulties involved in such analysis.

Complementarity and Wider Welfare Measures

12. The panel agreed that while adjusting the measured output of public services for complementarity might be inappropriate for the purposes of the national accounts, the issue underlined the case for developing better measures of economic welfare than GDP.
13. UKCeMGA was urged to develop the underlying methodological framework for such measures more rigorously and to document clearly the reasons for considering such an approach. One important consideration would be the treatment of equity issues.

14. The panel agreed that UKCeMGA's productivity articles would be an appropriate place in which to present such alternative welfare measures. This reflects the fact that the productivity articles have always been intended to provide the public with information on which to make judgements on value for money, as well as to describe potential measures for inclusion in national accounts.
15. It is possible that that 1.5 per cent earnings adjustment may turn out to be appropriate in a welfare measure parallel to Gross Domestic Income. But the series that include the adjustment should nonetheless be removed from the productivity articles for the time being, while such measures are being developed.
16. The panel also felt that there was a strong case for developing the use of satellite accounts to capture additional valuations of public services, although development of appropriate values and weights is probably some way off. Nevertheless, progress has been made in recent years in placing monetary values on non-market goods such as transport safety and certain environmental benefits. The extension of such values to other sectors would be very helpful in two respects. Within particular sectors, it would contribute to judgments about value for money. At the same time, it would provide additional guidance about the efficient allocation of resources across different parts of the public sector. But the panel noted that it might well be easier to identify changes in the value of public services over time than a robust measure of the absolute level at any given moment.
17. To ensure credibility when presenting alternative measures, it would be important to be clear and robust in defining and differentiating between the analytical frameworks appropriate for the national accounts and for broader welfare measures. The treatment of the public sector in any welfare measure would need to be consistent with treatment of the private sector.

Key Recommendations

To summarise, the panel recommended that:

1. Where there is ambiguity as to the distinction between price and output effects, the presumption should be that they are relative price effects rather than changes in volume. Adjustments based on public/private sector complementarity should only be made where there is a compelling argument in their favour.
2. Measures of public service output incorporating the 1.5 per cent earnings adjustment should be dropped from UKCeMGA productivity articles.
3. UKCeMGA should consider adjusting the output of the protective services to reflect changes in the volume of assets protected
4. ONS should develop an analytical framework for welfare measures broader than GDP, for which adjustments for private/public sector complementarity specific to particular services may be appropriate.
5. Welfare measures for particular services should be published in the productivity articles, once the analytical framework is fully developed and explained.