



**Atkinson Review:  
Interim report**

Measurement of Government  
Output and Productivity for  
the National Accounts

July 2004

# **Atkinson Review: Interim report**

---

Measurement of Government Output and  
Productivity for the National Accounts

© Crown copyright 2004

Published with the permission of the Controller of Her Majesty's Stationery Office (HMSO)

ISBN 0 11 703447 9

Applications for reproduction should be submitted to HMSO under HMSO's Class Licence:  
[www.clickanduse.hmso.gov.uk](http://www.clickanduse.hmso.gov.uk)

Alternatively applications can be made in writing to:

HMSO Licensing Division

St Clement's House

2-16 Colegate

Norwich NR3 1BQ

# Contents

	Page numbers
Preface	vi
<b>1. Introduction</b>	
Terms of Reference	1
The Review Team	2
Consultations	2
Actions and Action Plans	3
Contents of the Interim Report	3
<b>2. Measuring Government Output in the UK</b>	<b>5</b>
Conventional Approach to Measuring Government Output	5
Developments over Time in ONS (and CSO) Practice	6
Technological and Institutional Change in the Public Sector	10
Measures of Government Output, Inputs and Productivity: An Overview of the Principal Estimates	11
Findings for Individual Spending Functions	14
Data Sources	19
Conclusions	25
<b>3. The International Context</b>	<b>26</b>
International Guidelines: UN System of National Accounts and OECD Productivity Manual	26
European System of Accounts (ESA 95) and the Eurostat Handbook on Price and Volume Measures in National Accounts	30
Progress of Other Countries in Measuring Government Output	32
Summary	34
<b>4. The Functions and Limits of National Accounts</b>	<b>35</b>
Macroeconomic Management	35
Assessing Overall Economic Performance and Welfare	37
Relation with Government Performance Measures	38
International Comparisons and Obligations	39
Micro-Macro Links	40
The Limits of National Accounts: Conventions and Margins of Error	41
Conclusions	42

	Page numbers
<b>5. A Methodological Framework</b>	<b>43</b>
The Concept of Added Value	43
Parallel with the Private Sector	44
The Coverage of Output Indicators	46
Inputs and Productivity	47
Triangulation	50
Quality Change and the Complementarity between Public and Private Output	52
Summary of Principles	55
<b>6. Implementation</b>	<b>57</b>
Output Measures	57
Input Measures and Deflators	59
Further Consideration	61
Process of Implementation	62
Publication and Presentation of Estimates	63
Resource Costs of Statistics	65
<b>7. Health: A Progress Report</b>	
Introduction	<b>66</b>
How UK Health Outputs are Currently Measured	67
International Context	68
Critique of Current Health Output Methods	69
Way Forward on Output Methods	73
Inputs and Deflators	75
Triangulation and Productivity Measurement	78
Summary of Work in Progress and Future Directions	79
<b>8. Education: A Progress Report</b>	
Introduction	<b>81</b>
How UK Education Output is Currently Measured	83
International Context	85
Critique of Current Education Outputs	87
Ways Forward on Output Methods	88
Inputs and Deflators	97
Triangulation and Productivity Measurement	99
Summary of Work in Progress and Future Directions	99
<b>9. Public Order and Safety: A Progress Report</b>	
Introduction	<b>101</b>
How UK Public Order and Safety Output is Currently Measured	103
International Context	105
Critique of Current Public Order and Safety Output Methods	106
Way Forward on Output Methods	109

	<b>Page numbers</b>
Inputs and Deflators	111
Triangulation and Productivity Measurement	112
Summary of Work in Progress and Future Directions	112
<b>10. Social Protection: A Progress Report</b>	
Introduction	<b>114</b>
How UK Social Protection Output is Currently Measured	118
International Context	120
Critique of Current Social Protection Output Methods	121
Way Forward on Output Methods	123
Inputs and Deflators	125
Triangulation and Productivity Measurement	125
Summary of Work in Progress and Future Directions	126
<b>11. Summary of Main Conclusions</b>	<b>127</b>
<b>Appendices</b>	
A UK Consultation	134
B International Consultation	135
C International Guidance	147
D How OECD Countries Measure Government Output	157
<b>Glossary</b>	<b>165</b>
<b>Abbreviations</b>	<b>168</b>
<b>References</b>	<b>170</b>

# Preface

---

The National Statistician asked me in December 2003 to conduct an independent review of the measurement of government output in the National Accounts, and to produce an interim report by July 2004. Both the task and the timetable were challenging. The fact that it has been possible to produce this interim report is due to the remarkable quality of the team assembled for the review, headed by Joe Grice, with Aileen Simkins as co-director. I would like to thank ONS, the Treasury, departments, and the Bank of England for seconding staff to the review.

While I take full responsibility for the limitations of this report, and for the views expressed, I should stress that its production has been very much a team effort. I am most grateful to all members of the review team for their contributions over the past six months.

**Tony Atkinson**

July 2004

# 1 Introduction

---

## Terms of Reference

**1.1** The terms of reference of the review set out by the National Statistician were:

‘To advance methodologies for the measurement of government output, productivity and associated price indices in the context of the National Accounts, recognising:

- a) the full scope of government outputs;
- b) differences in the nature and quality of these outputs over time;
- c) the relationship between government outputs and social outcomes;
- d) the need for comparability with measures of private sector services’ output and costs;
- e) the existing work of the Office for National Statistics (ONS); and
- f) the appropriate measurement of inputs, including quality and the distinction between resource and capital, so that, together with the measurement of output, light can be thrown on developments in government productivity.’

These terms of reference have set the framework for the review.

**1.2** Two aspects, in particular, have shaped the approach adopted. The first is that our brief is to examine the measurement of government output within the context of National Accounts. The end result should be reflected in the National Accounts or associated statistics. While we are mindful of the performance indicators used in Public Service Agreement (PSA) targets and other management agreements, and refer to these at a number of junctures, these are not our concern in this review. We are seeking aggregate indicators that can form part of the National Accounts, not complete tool kits for the management and audit of government activities.

**1.3** The second aspect is the three-way association between government inputs, outputs, and productivity. While ONS has published figures on government productivity (*Economic Trends*, July 2003), productivity at present appears only implicitly in the National Accounts. We have taken as one of our central concerns the establishment of a clear and transparent relationship between inputs, outputs and productivity. The measurement of government productivity, in an aggregate way, is at the heart of our inquiry.

- 1.4** The aim of the review is to establish the future strategic direction for work in this area. When completed, in our final report, the work will, we hope, form part of the development programme for National Accounts in the United Kingdom. More widely we hope it will influence the future shape of the United Nations *System of National Accounts* (SNA) and the *European System of Accounts* (ESA).

### The Review Team

- 1.5** The review has been conducted by Sir Tony Atkinson, Warden of Nuffield College, Oxford, with the support of the review team based in ONS. The team has been headed by Joe Grice, Deputy Head of the Government Economic Service, on secondment from the Treasury with Aileen Simkins from the Department of Health as co-director. It has consisted of ONS staff and staff seconded from the Bank of England, the Treasury and government departments responsible for public services. The other members of the team to date are:

Alwyn Pritchard, Phillip Lee, Nicola Mai, James Hemingway, Janet Snelling, Sukwinder Mehmi, Amanda Tuke, Michael Lyon, Ben King and Liz de Freitas.

- 1.6** The review has benefited greatly from the help of a Co-ordination Group, chaired by Len Cook, the National Statistician. The list of members is in Appendix A. We are grateful to members for their views and for facilitating work within their departments.

### Consultations

- 1.7** In the review, we have sought to consult widely, within the limitations imposed by the tight timetable. The review organised an open consultation meeting at the Bank of England in May 2004, to which a wide variety of interested parties were invited. Sir Tony Atkinson had meetings with the Cabinet Secretary, the Monetary Policy Committee of the Bank of England, the Statistics Commission, the Chair of the Audit Commission, the Comptroller & Auditor-General, and the Head of the Prime Minister's Delivery Unit. Members of the review team have made presentations at conferences and seminars, from which we have received useful comments. We had a valuable meeting with representatives of the Royal Statistical Society.
- 1.8** In making its recommendations, the review team has had as its objective that the resulting series should meet the criteria for designation as National Statistics. All National Statistics are produced to the professional standards set out in the National Statistics Code of Practice. The statistical methods must meet the demands of peer review. To this end, the review team has consulted extensively with other branches of ONS, in particular with the National Accounts Group.
- 1.9** The review team consulted the devolved administrations, visiting Belfast, Cardiff and Edinburgh.

**1.10** The review team paid particular attention to international consultation, in view of the importance of international guidelines and because it wished to learn from the experience of other countries. The team visited Eurostat and the Organisation for Economic Co-operation and Development (OECD). The team visited the Statistics Offices of Finland, Italy, the Netherlands, Norway and Sweden – all of whom are in the vanguard on this particular agenda. Appendix B gives details of international contacts. We are most grateful for the time and welcome given to us. The team had a very helpful discussion with Rob Edwards, Australian Deputy Statistician, during his visit to ONS in March 2004, as well as with Jack Triplett, formerly of the US Bureau of Economic Analysis. None of these bodies should of course be held in any way responsible for the views expressed in this report.

### Actions and Action Plans

**1.11** Our starting point has been the existing work of ONS, and our conclusions build on the substantial progress that has been made by ONS since it began in 1998 to measure government output directly. The field that we are reviewing is a dynamic one, and the report needs to be read bearing in mind that advances are continually being made in ONS practice. Members of the review team have, for instance, been involved in the development of improved measures of Health output introduced by ONS in *Blue Book* 2004.

**1.12** Progress towards improved measures of government inputs and outputs depends very much on co-operation between government departments, the Treasury, and ONS. To facilitate this, action plans have been set up with key service delivery departments (Department of Health, Department for Education and Skills, Department of Work and Pensions, Home Office, Office of the Deputy Prime Minister) and are in the process of being set up with those departments responsible for supplying expenditure data (such as the Treasury) to ONS. We should like to stress the essential role played by this co-operation and urge that departments continue to accord priority to the provision of data on a reliable and timely basis. We hope that our reports will underscore the importance of this activity.

### Contents of the Interim Report

**1.13** This report is preliminary in two senses. First, a number of the ideas put forward are intended to engender further discussion and do not represent our final judgements. The report will, we hope, give rise to public debate and comments that we can take into account in preparing the final report, due in January 2005. Secondly, the implementation of new measures of government output and productivity takes time. A great deal of work is necessary for their construction and for their evaluation, including the rigorous ONS process of peer review. This work is under way but not yet complete in most cases. It is not therefore possible to reach definite conclusions in this interim report about the desirability and feasibility of all the new approaches discussed here.

**1.14** The report is about methodology. We have reached a number of interim conclusions, and these are collected in Chapter 11. These conclusions, if adopted by ONS, may lead to changes in measured government output and the implied indicators of government productivity. But this report does not contain any new substantive findings with regard to the figures for government output or productivity. The publication of National Accounts statistics is the responsibility of ONS, and any new measures will be published by them, as with the Health output figures referred to in paragraph 1.11.

# 2 Measuring Government Output in the UK

- 2.1** In this review, we are concerned with the measurement of the volume of government output relative to the volume of government inputs, and with the implied measure of government productivity. These measures are important because the functioning of public services are a matter of widespread public interest. These measures are significant at a macroeconomic level, because government output represents a sizeable part of Gross Domestic Product (GDP). In broad terms, a one per cent per year faster growth rate of government output raises the overall GDP growth rate by some 0.2 per cent.
- 2.2** For the purpose of this review, the term ‘government’ should be read broadly, to cover all those agencies which provide public services. Thus, for example, the National Health Service (NHS) and local authority provision of social services within our purview. The precise boundaries of the government sector are important, but are not particularly germane to our report. At a number of points, it will, however be important to distinguish between *individual* services (those consumed by individual households) and *collective* services provided to society as a whole. Government output is largely non-market output in the sense that it is supplied free or at prices that are not economically significant.
- 2.3** Government features as an identifiable component in two different measures of GDP. First, there is the final consumption expenditure of government, a component part of GDP (Expenditure). For 2002, the *Blue Book 2003* showed *individual* government final consumption expenditure of £129bn, and *collective* government final consumption expenditure of £80bn at current prices. Total GDP at current market prices in 2002 was £1,044bn. Here our concern is primarily with a second measure: GDP at constant prices. This volume measure provides an indicator of the growth rate of the economy. Government output, at constant prices, enters into this GDP measure. Individual volume measures for Health, Recreation and Culture, Education, Social Protection and Housing are also published in the *Blue Book*.

## Conventional Approach to Measuring Government Output

- 2.4** In many countries, and in the United Kingdom from the early 1960s to 1998, the output of the government sector has been measured by convention as of value equal to the total value of the inputs; by extension the volume of output has been measured by the volume of inputs. This convention regarding the volume of government output is referred to below as the (output = input) convention. The inputs taken into account in recent years in the United Kingdom are the compensation of employees, the cost of goods and services and a charge for the consumption of fixed capital. In earlier years, and in other countries, including the United States, the inputs were limited to employment.

**2.5** Wide use of the convention that (output = input) reflects the difficulties in making alternative estimates. Simply stated, there are two major problems: (a) in the case of collective services such as Defence or Public Administration, it is hard to identify the exact nature of the output, and (b) in the case of services supplied to individuals, such as Health or Education, it is hard to place a value on these services, as there is no market transaction.

**2.6** The rationale for the convention adopted in the United Kingdom for many years was described by Rita Maurice in her 1968 edition of *National Accounts: Sources and Methods*:

‘There are inherent difficulties in measuring, at constant prices, the output of government services, the exact nature of which cannot be precisely determined. ... In general, the solution adopted has been to use changes in numbers employed as an indicator of changes in output. The implied assumption of no change in productivity is obviously not satisfactory since it takes no account of the use of more modern equipment, such as computers. The problem has been discussed at a number of international conferences without any generally acceptable solution being reached. In some countries the assumption has been made that productivity in the public services increases at the same rate as productivity in all other industries, or in all other services ... in the United Kingdom accounts the assumption of no change in productivity has been preferred’ (Maurice, 1968, pp 44–45).

**2.7** The criticism that this (output = input) convention neglects changes in productivity is clearly well founded, and has been articulated many times by ONS: for example, in *Economic Trends*, February 1998. To the extent that productivity grows, the growth rate of government output is understated, and hence the overall growth rate of GDP is understated. Countries with larger public sectors, such as European economies, will on this account have lower relative growth rates, other things equal, than countries with smaller public sectors, such as the United States. Where, as in Germany, an assumed productivity growth factor has been added in estimating the contribution of government output to GDP, the growth rate will appear higher than in countries which assume that there is no productivity growth. International comparisons are thus affected.

### Developments Over Time in ONS (and CSO) Practice

**2.8** ONS has consistently taken the view that the output of the public sector should be included in GDP, rejecting any suggestion that it should be treated as a purely intermediate output. This is a view that we fully share, not least because it is adopted by the United Nations *System of National Accounts* (SNA) and by the *European System of Accounts* (ESA 95), which is based on the SNA 1993. (The international guidelines are discussed in Chapter 3.)

- 2.9** The SNA has, however, evolved in its approach to measurement of government output. The new approach introduced in the United Kingdom with the 1998 *Blue Book*, with its emphasis on actual personal consumption, opened the way to the direct measurement of the output of the government sector. ONS was in the vanguard of national statistical offices in initiating a project to introduce new measures of output for public sector non-market production. The project was described by Henry Neuburger and David Caplan in *Economic Trends*, February 1998, and its progress can be traced in a series of articles: David Caplan, *Economic Trends*, October 1998, Michael Baxter, *Economic Trends*, September 2000, Alwyn Pritchard, *Economic Trends*, May 2001, Timi Ashaye, *Economic Trends*, November 2001, Alwyn Pritchard, *Economic Trends*, May 2002, and Alwyn Pritchard, *Economic Trends*, July 2003.
- 2.10** Since 1998, ONS has moved increasingly towards the replacement of the (output = input) approach by direct measures of the volume of government output. The major progress made to date is summarised in Table 2.1. Successively, different sectors have been converted to a direct measure of output, beginning with Health, Education and the Administration of Social Security. The direct estimates now cover some two-thirds of General Government Final Consumption, which is an impressive achievement.

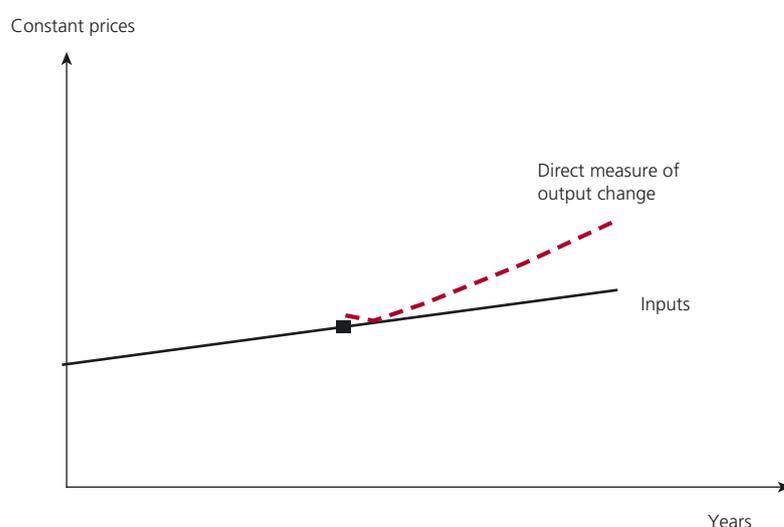
Table 2.1 Post-1998 Developments in ONS Measurement of Government Output

Function	Per cent government spending in 2000	Date introduced	Main components	Data source coverage	Reference
Health	30.3	Introduced in 1998 <i>Blue Book</i> , with data from 1986, method updated 2004	Hospital Cost Weighted Activity Index, Family Health Services (number of GP consultations, etc)	Mainly data for England	<i>Economic Trends</i> , October 1998
Education	17.1	Introduced in 1998 <i>Blue Book</i> , with data from 1986	Primary and secondary schools pupil numbers – Quality adjustment of 0.25 per cent	UK figure for pupil numbers in nurseries and primary and secondary schools	<i>Economic Trends</i> , October 1998
Administration of Social Security	2.7	Introduced in 1998 <i>Blue Book</i> , with data from 1986	Number of benefit claims for 12 largest benefits – No allowance for collection of contributions	United Kingdom	<i>Economic Trends</i> , October 1998
Administration of Justice	3.0	Partially introduced in 2000 <i>Blue Book</i> , with full impact in 2001 <i>Blue Book</i> , with data back to 1994 Q1	Number of prisoners, legal aid cases, court cases, and probation cost-weighted activity index	Great Britain for prisons, otherwise England and Wales	<i>Economic Trends</i> , September 2000, for probation, November 2001
Fire	1.1	Introduced in 2001 <i>Blue Book</i> , with data from 1994 Q1	Number of fires, fire prevention and special services	England, Wales and Northern Ireland	<i>Economic Trends</i> , November 2001
Personal Social Services	7.4	Introduced in 2001 <i>Blue Book</i> , with data from 1994 Q1	Children and adults in care and provision of home helps	England	<i>Economic Trends</i> , November 2001
Police	5.8	Experimental	Cleared-up crimes of different types	England, Wales and Northern Ireland	<i>Economic Trends</i> , May 2002

Source: ONS

**2.11** In these new estimates, ONS has taken the output, calculated on the basis of inputs, at a reference year, and the direct measures have been used in the form of changes since that date. It is important to understand what these measures do and do not do. The direct measures are of *changes in output*. They are not attaching a monetary value to the *level* of output recorded using the new indicator. The change is shown schematically in Figure 2.1. Starting from a particular base date, shown by the square, the direct output indicators, such as pupil numbers, are used to estimate the growth of output volume. They are not used to estimate the relation between inputs and outputs at the base date. Nor is any attempt made to calculate a current price measure that could be included in the measure of GDP at market prices.

Figure 2.1 **Schematic Representation of ONS Direct Output**



**2.12** The move away from the (output = input) convention has introduced a divergence in method between constant price and current price measures of GDP. This may be seen from the fact that the convention that the value of output was equal to the value of inputs meant that there was no net operating surplus or deficit to be attributed to the public sector output. If, as shown in Figure 2.1, output exceeds inputs, there is an operating surplus; the same would arise if there were an assumed productivity growth (as in the German estimates).

**2.13** Although not explicitly recognised, the post-1998 approach of ONS has affinities with that used in the United Kingdom in the 1950s and early 1960s. As argued by the Central Statistical Office (CSO) (the predecessor of ONS) in 1956, ‘even a crude measure of [government] output is assumed to be preferable to an index based on total cost, at constant prices, of the factors of production ... the latter method is only used *faute de mieux*, e.g. in public administration, where it is difficult to conceive of any sensible measure of service rendered.’ (CSO, 1956, pp 42 and 43). ‘Thus the administration of national insurance is represented by the numbers insured and the numbers in receipt of benefit; the hospital services by the number of patients and the number of hospital staff; courts of justice by the number of cases tried.’ (CSO, 1956, pp 42).

- 2.14** According to Maurice, writing in 1968, the direct indicators ‘proved unsatisfactory and have been abandoned’ (1968, pp 87). The reasons she gives are ‘the difficulty of finding indicators’ and ‘the need to keep the output estimates consistent with the expenditure estimates’ (1968, pp 87). According to Levitt and Joyce, the direct output measures used in the 1950s and early 1960s were ‘heavily criticised’ (1987, pp 50) on the grounds of the arbitrary nature of the weights employed (for example, for teachers and pupils), the tenuous nature of the link with outputs, and the issue of quality change. They cite the example of the impact of government policies designed to reduce class sizes. Since the output indicator for Education gave half weight to pupils and half to teachers, output per teacher was assumed to be a declining function of the number of teachers; hence, a rise in the number of teachers increased output less than proportionately. It is not, however, evident that this is inferior to the assumption that labour productivity is independent of the number of pupils per teacher. Certainly, the estimates of Beales (1967) showed that between 1959 and 1963 the direct method led to estimates of output increase that were lower than those obtained using the (output = input) convention. For Education, the direct method yielded estimates of 13.5 per cent increase in output, compared with 19.4 per cent, and of 6.2 per cent for Health, compared with 9.2 per cent (figures cited by Levitt and Joyce, 1987, pp 50) Beales noted that this would not always be the case.
- 2.15** From this earlier experience, we draw two main conclusions. The first is that the design of direct output measures needs considerable care. It is not necessarily the case that ‘even a crude measure of [government] output is ... preferable to an index based on total cost’. The fact that it is not easy to obtain direct indicators means that better measures are likely to require significant investment of resources. Direct measures of output need to be continuously monitored to ensure that they are capturing changes in quality. The second conclusion is that ONS has to steer a careful course with regard to changes in government policy, guaranteeing the independence of the approach to measuring government output while ensuring that its implementation reflects the realities of public spending and circumstances.

### **Technological and Institutional Change in the Public Sector**

- 2.16** There is a strong case for devoting significant resources at this time to improving the measurement of public sector output on account of its increased saliency in policy-making and public debate. Interest in the measurement of government output has become much more intense in recent years as a result of a number of developments in the public sector, reflecting changed government priorities, the concerns of citizens and voters, institutional change, and the new possibilities opened by technological developments. As it was put by the National Statistician when he announced the review, ‘the amount of resources allocated to public services has increased. Delivery and management mechanisms have developed and are more complex. There is an increasing emphasis on the quality of service for the customer. As a result there are greater demands on, and expectations of, measures of government output’.

- 2.17** The public sector is changing rapidly. Functions have been transferred between government departments: for example, the transfer of the levying of National Insurance contributions, and the payment of Child Benefit, from the Department of Work and Pensions to the Inland Revenue. Under successive governments, there have been major policy initiatives directed at changing methods of working and these have often involved shifting institutional boundaries. New delivery mechanisms have been introduced, such as NHS Direct. Policy changes have involved transfers between functions, such as residential or domiciliary care, provided by personal social services, replacing hospital days. An increasing part of service provision may involve sourcing from the private sector.
- 2.18** Ideally, the measure of public sector output should be invariant with respect to changes in the organisation of delivery, but this may not be easy to achieve in statistical practice. Moreover, we have been impressed by the fact that the move to direct output measurement may have made this problem more difficult. With the (output = input) convention, institutional change may not affect the adding up of input outlays within a department, or may involve a relatively simple transfer of input spending. With direct output indicators, there is no guarantee that the effects on the separate output indicators within a department will cancel; where functions are transferred across departments, new output indicators may need to be constructed.
- 2.19** The public sector is also affected by rapid technological change. This may take the form of sector-specific change, as in medicine, or of the general impact of information and communications technology (ICT). Technological change is not confined to the public sector and its impact has been much discussed with regard to the private business economy. In this sense, we may be able to learn from the private sector experience. However, extensive debate in the United States has shown that the measurement of the output of the private service sector is itself a challenging problem. For many years, there was concern that computers were present everywhere except in the productivity statistics (a paradox enunciated by Robert Solow). The impact of technological change may appear only with a delay, and the benefits of ICT may appear in quality change that is not recorded.

### **Measures of Government Output, Inputs and Productivity: An Overview of the Principal Estimates**

- 2.20** We believe that it is useful to begin by ‘looking at the numbers’. In this and the next section, we examine the most recent figures available when we started work in January 2004: those published in the July 2003 *Economic Trends* article. (We have not updated the figures since our concern here is mainly with the methodology.)
- 2.21** The figures for general government are summarised in Table 2.2 and illustrated in Figure 2.2, volume of output takes the 1995 volume of inputs figure as a starting point, so that the measure of productivity is of the change relative to that date. As may be seen from the figures in Table 2.2, the volume of output rose over the period, though at a rather faster rate after 1998. The volume of inputs (in constant prices) increased only marginally until 1998 but then rose sharply. The picture is shown in Figure 2.2, with the output and input lines crossing in 1999.

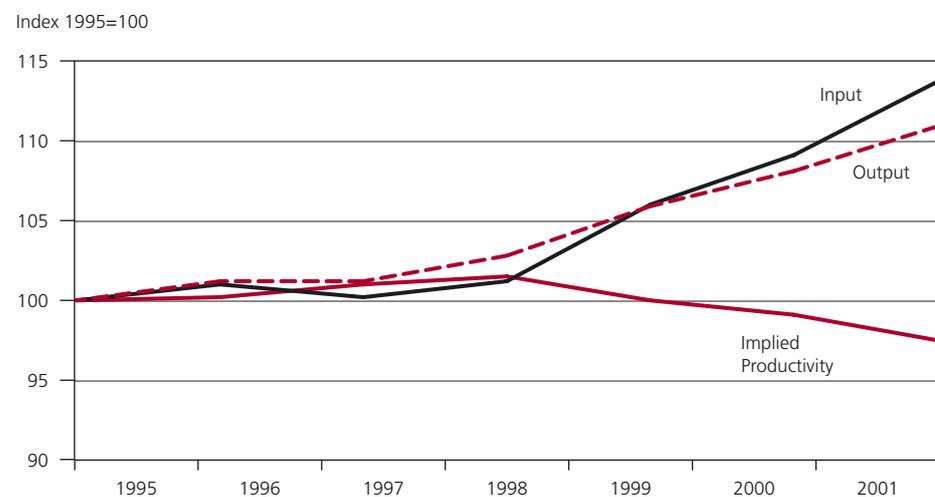
Table 2.2 Recent Estimates of Government Output, Inputs and Implied Productivity

£ million	1995	1996	1997	1998	1999	2000	2001
Volume of government output at 1995 prices	141,031	142,702	142,779	144,991	149,419	152,524	156,361
Volume of government output, index (1995=100)	100.0	101.2	101.2	102.8	106.0	108.2	110.9
Volume of government inputs at 1995 prices	141,031	142,388	141,371	142,785	149,441	153,877	160,320
Volume of government inputs, index (1995=100)	100.0	101.0	100.2	101.2	106.0	109.1	113.7
Implied index of productivity (1995=100)	100.0	100.2	101.0	101.5	100.0	99.1	97.5

Source: *Economic Trends, July 2003, Table 2*

**2.22** Use of direct measures of output means that we now have an implied measure of productivity. Since the output series moved less sharply, the implied productivity series (see Figure 2.2) first rose, by some 1.5 per cent, and then fell from 1998 to 2001 by some 3.9 per cent. This, it should be stressed, is an *implied* measure of productivity, and it may or may not be consistent with independent evidence on the productivity performance of the public sector.

Figure 2.2 Total Government Output, Input and Implied Productivity



Source: *Economic Trends, July 2003, Table 2*

- 2.23** It may be helpful to set out the key ingredients, with the amounts in 1995 and 2001:
- **Expenditure at current prices on the inputs purchased by government** to produce its outputs, usually referred to as General Government Final Consumption Expenditure at Current Prices: £141bn in 1995 and £192bn in 2001, an increase of 36 per cent;
  - **Volume of government inputs, adjusted for the increases in the prices of inputs**, using appropriate price indices: at 1995 prices, £141bn in 1995 and £160bn in 2001, implying that on average input prices had risen by some 20 per cent and that the volume of inputs had increased by some 13.5 per cent;
  - **Volume of government output** at 1995 prices: £141bn in 1995 and £156bn in 2001, a rise of 10.6 per cent.
- 2.24** On occasion, the first series is divided by the third to obtain the ‘implied deflator’ or cost per unit of government output. This deflator includes, however, both price and productivity elements. It is equal to the input price increase divided by the productivity increase. It seems to us preferable to separate the two elements.
- 2.25** In the three definitions above, we have set in bold the key elements. Where output continues to be measured by inputs – the conventional method – only the first two elements are involved. This applies to Defence, General Public Services, Economic Affairs, Environmental Protection, Recreation and Culture, Housing and Community Amenities.
- 2.26** For the remainder of public spending, around two thirds, there is an independent measure of output, and hence productivity. (In the case of police spending, the output measure is experimental and not yet included in the National Accounts but this function is included here with those directly measured).
- 2.27** It is possible that the productivity indicator is capturing a real phenomenon. We earlier motivated the change to a direct output measure by the consideration that the (output = input) convention ignored productivity change, but there is no necessary reason to expect productivity to increase. Many public services, like social services, involve an essential human input. In the caring professions, there may be little scope to increase labour productivity. (Although at the same time, the worker may be aided by capital, such as use of a car rather than public transport, allowing more clients to be visited in a day. Computerisation may allow care workers to be allocated more efficiently.)
- 2.28** The expansion of public services may be subject to diminishing returns. Historically, initial reductions in neo-natal mortality may have been relatively easy to achieve, but driving down rates today below their current level may involve much more expensive interventions. In current terms, when schools have reduced truancy by ten per cent, they may find it increasingly difficult to achieve further reductions. When re-offending has been reduced by five per cent, the next five per cent may mean working with more difficult cases. Diminishing returns do not, of course, mean that the expansion is unjustified. The value of the additional output may still exceed the cost of the inputs. As already noted, the present estimates are silent on this relationship: they tell us nothing about the level of productivity, but rather the rate of change.

**2.29** Finally, it is possible that there are lags between inputs and outputs, and that a marked increase in public spending, such as that which has taken place in recent years, may show up in improved output indicators only at a later date.

**2.30** For all these reasons, the estimates may be correctly measuring the change in productivity (although they tell us nothing about the level). But, equally, it is possible that the estimates are in error and are failing to capture important elements. The accuracy of the productivity figure depends on the accuracy of the three constituent elements. The productivity increase will be **understated** if in the terminal year:

- The direct output measure is too low;
- The measure of expenditure on inputs is too high;
- The price index for inputs is too low.

The productivity increase will, on the other hand, be **overstated** if in the terminal year

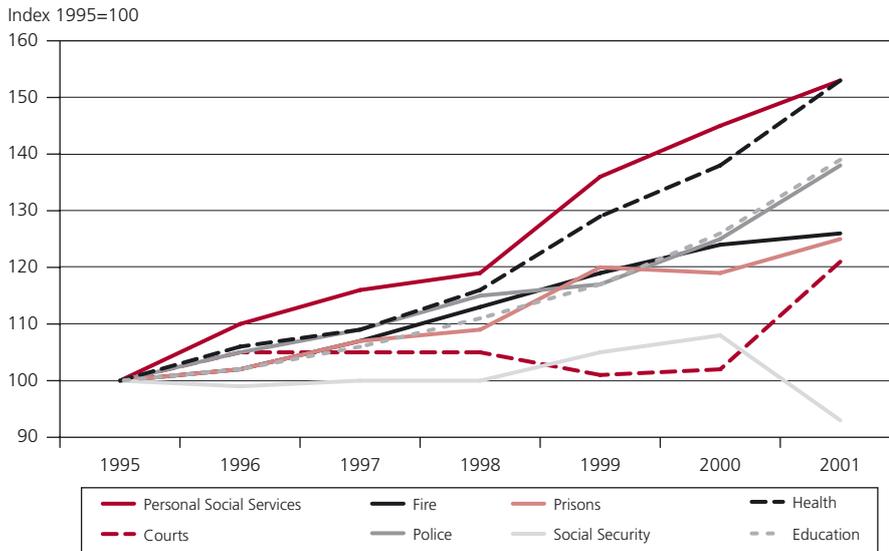
- The direct output measure is too high;
- The measure of expenditure on inputs is too low;
- The price index for inputs is too high.

### Findings for Individual Spending Functions

**2.31** We now consider separately the eight functions for which there are direct measures of output. Of these, Health is the largest, accounting for 42 per cent of the total, followed by Education, accounting for 28 per cent. Next come Personal Social Services (ten per cent) and Police (eight per cent). As noted above, the output measure for Police is experimental and not yet included in the National Accounts. Four of the eight functions therefore account for nearly 90 per cent of the total public spending where there is direct output measurement.

**2.32** In order to compare them, we begin by plotting in Figure 2.3 the total expenditure (1995=100). The four largest functions all saw the largest increases in spending. Health and 'Personal social services' expenditure rose 53 per cent from 1995 to 2001. Education and Police both saw rises of approximately 38 per cent. Prisons and Fire increased by 25 per cent, followed by Courts with a 22 per cent increase. Administration of Social Security saw a reduction of seven per cent. The timepath is not steady; as is well known, spending accelerated post-1998. In the case of Health, for instance, there was a deceleration from 1995 to 1998, followed by a sharp upturn. The expenditure on Courts exhibited a variable pattern over time.

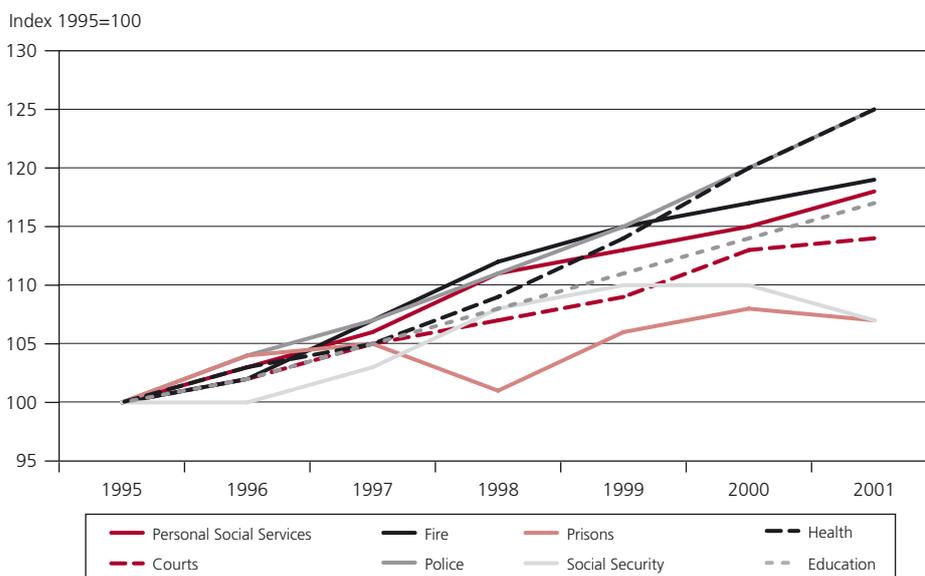
Figure 2.3 Expenditure by Function



Source: Economic Trends, July 2003, Table 3

**2.33** In order to translate these into real increases in inputs, we need to allow for the increases in the cost of inputs. Figure 2.4 shows that these differed considerably across the eight functions. The largest rises (25 per cent between 1995 and 2001) are shown by the cost indices for Police and Health, but even for these the time path is different, with the Police price index rising faster in the early part of the period. The next group, with a rise of around 17.5 per cent, are Fire, Personal Social Services and Education, followed closely by Courts. Finally, the increases in input prices were only about seven per cent for Prisons and Administration of Social Security. The extent of these differences by function is rather surprising. The reader is prompted to ask about the provenance of the price deflators.

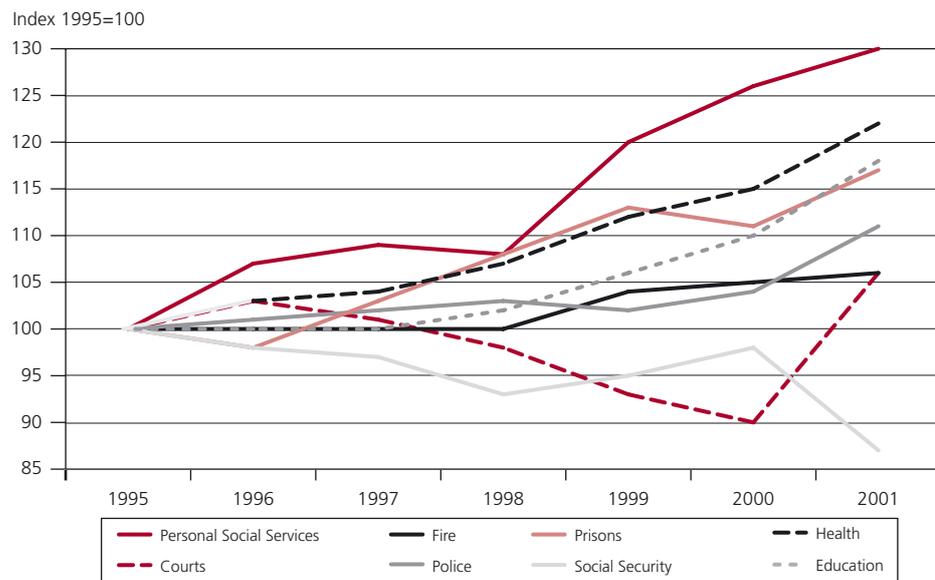
Figure 2.4 Overall Price Indices



Source: Economic Trends, July 2003, Table 3

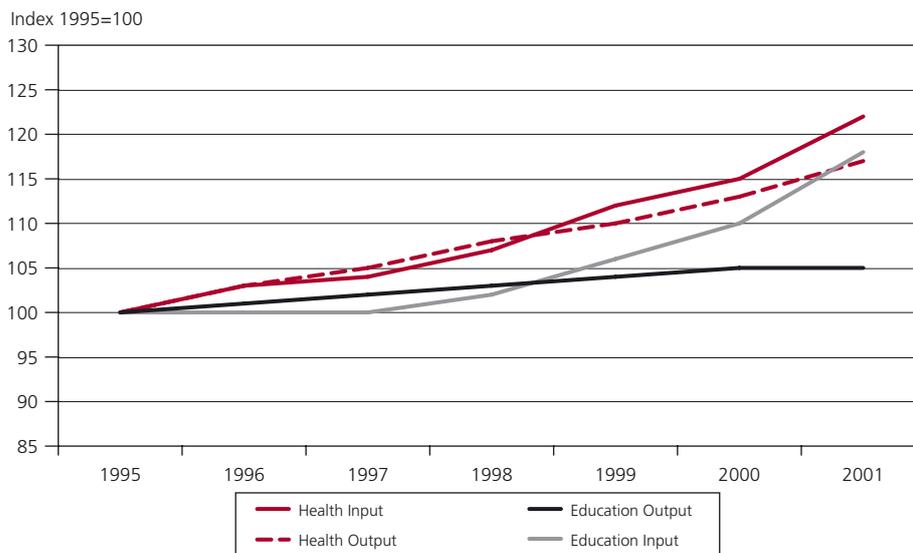
**2.34** The result of these differential movements in input prices is that the input volume indicators look rather different from the current expenditure series. If we just concentrate on the four largest functions, then Health and Personal Social Services had grown by the same amount in terms of spending, but Personal Social Services has now a significantly larger volume increase, as input prices are measured as rising less. The same applies to Police and Education spending: both rose by the same money amount, but the larger input price rise for police implies a smaller volume increase. The input volumes are shown in Figure 2.5.

Figure 2.5 **Volume of Inputs**



**2.35** The final part of the story is the direct measure of output, shown, together with the input volume series, for the two largest functions, Health and Education in Figure 2.6. In both cases, the output estimates for 2001 are below the input series. For both functions, productivity measured in this way is lower than in 1995. Although ONS caution that ‘the data are not yet robust enough to report the productivity calculations for individual government functions’ (*Economic Trends*, July 2003, pp 32), the phenomenon described in aggregate in the previous section applies to both the largest two spending functions.

Figure 2.6 Direct Estimates of Output and Input Volumes



Source: *Economic Trends*, July 2003, Table 3, 6 and 8

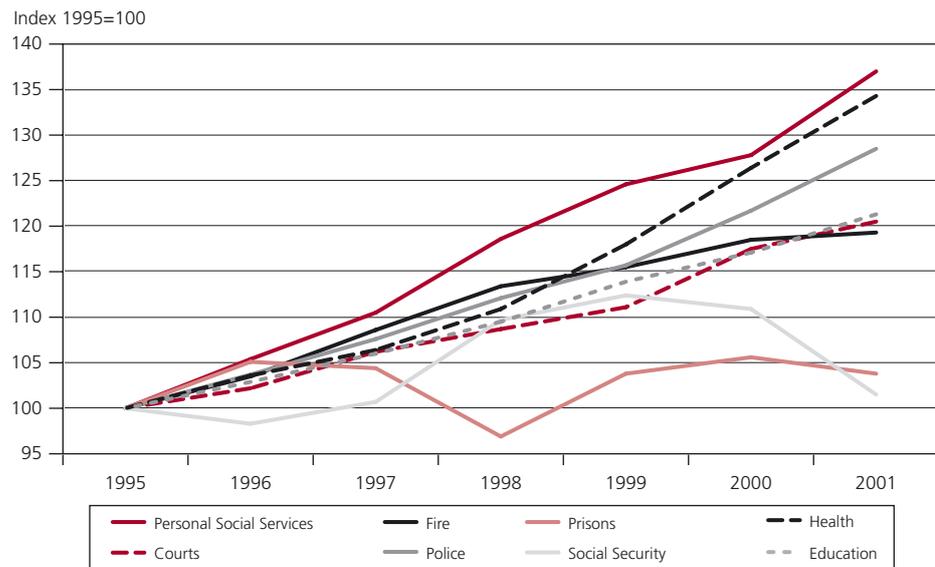
## Inputs

- 2.36** The changes made to the output measurement methodology mean that we have tended to focus attention on the output side, but it is equally important to investigate the measures of inputs.
- 2.37** The input side is best considered in terms of the three disaggregated input categories given in the National Accounts figures: labour, goods and services, and capital consumption. In the case of Education, for example, the labour input consists of the wages and salaries paid to teachers, school secretaries, caretakers, and other employees, together with the costs of employing them, such as National Insurance and pension contributions. Goods and services would include expenditure on exercise books, pens, lighting, heating, supply teachers, transport services, and items such as data processing services. Capital consumption is the physical depreciation of the stock of fixed assets, where these would include, in the case of Education, buildings, equipment, and computer hardware/software.
- 2.38** Labour is usually assumed to be the most important input for public services, and this is certainly the case for Prisons (64 per cent in 1995), Education (71 per cent), Police (87 per cent) and Fire (89 per cent). In the four other functions, labour is a smaller part of the total cost, reflecting changes in organisation noted above. In Personal Social Services, and the Administration of Social Security, labour input and the purchase of goods and services are not dissimilar in magnitude; in Courts, the share of goods and services reached 60 per cent (in 1995). In Health, almost all the expenditure is accounted for by goods and services (91 per cent in 1995). As is explained in *Economic Trends*, July 2003, pp 36, this arises because primary care trusts commission services from hospital trusts and other providers, where these are classified as either public corporations or private producers. The government purchases health care on behalf of patients but does not employ the staff who produce them. Capital consumption is six per cent in Prisons, four per cent in Education and in other areas 2.5 per cent or less.

**2.39** It should be noted that there was considerable change over the period in the composition of inputs. In July 2003, moreover, ONS announced a change in the sector classification of NHS Trusts, which from July 2004 will be shown as integral to government. This will increase the compensation of employees and reduce the level of government health procurement spending.

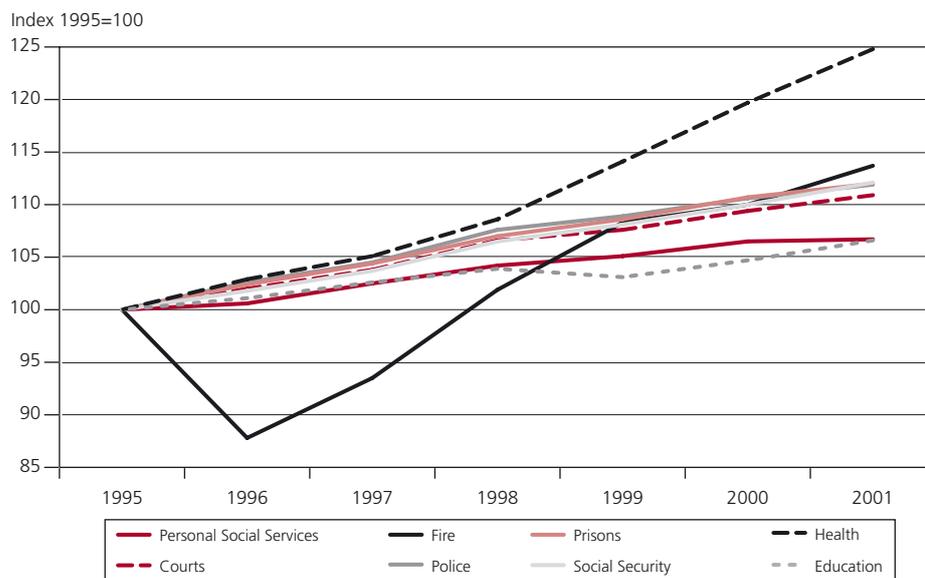
**2.40** Figures 2.7 and 2.8 show the price indices for labour and for goods and services. There are some surprising features. The wage cost index appears to have scarcely increased for Administration of Social Security and to have risen by only four per cent over seven years for Prisons. There is a much larger rise in the price index for goods and services for Health than for the other functions. Examination of these differential movements leads one to ask about their construction. The price index for capital consumption (not shown) also exhibits considerable variation across the functions for the period as a whole, after several years when the figures were identical for five of the functions. It is also not evident why the price index for capital consumption fell between 1996 and 1997 for Health and Personal Social Services, and for five of the eight functions between 1998 and 1999.

Figure 2.7 **Price Indices: Labour**



Source: *Economic Trends, July 2003, Table 3*

Figure 2.8 Price Indices: Goods and Services



**2.41** Consideration of the findings for individual functions suggests that all three ingredients – direct output measures, spending on inputs, and the deflation of inputs – warrant closer investigation. We turn now to the data sources.

### Data Sources

**2.42** This section provides a brief summary of the existing process of data collection and analysis. (Further detail by service area is provided in Chapters 7 to 10.) Note that this refers to the situation as at early 2004, and not necessarily to the *Economic Trends* article cited earlier.

**2.43** Table 2.3 shows the sources of data for the existing output volume measures, used to compile total General Government Final Consumption as chained volume measures. Where output volumes are measured directly (parts of Public Order and Safety, Health, Education and Social Protection), the coverage of the indicators in terms of countries of the United Kingdom varies, as is shown in the table.

**2.44** The table includes direct measures of output volume (shown in **Bold**) used in implied productivity calculations, and deflated input measures, used as measures of output volume (shown in *italics*). As before, the measures for Police, although only experimental, are shown as direct measures.

**2.45** It is evident from this table that the process of data assembly is one of considerable complexity. It involves central government departments, the Treasury, local authority data supplied via the Office for Deputy Prime Minister (ODPM), and the devolved administrations. The Home Office, for example, supplies data on Police spending via the Treasury; the local authorities supply data on the probation service via ODPM.

- 2.46** A key element is the classification of government expenditure. There is an effective collaboration between ONS and the Treasury on economic categories which is communicated to departments, and this will be further improved when the Treasury supply data consistent with the new Classification of the Functions of Government (COFOG) breakdown in 2004/2005. (COFOG is the UN classification of government activities for the purposes of National Accounts.) However, because of processing at each stage of the data supply chain, it is our impression that there are still issues in communicating to departments how their figures have been included in the National Accounts, and of the need to reconcile the classifications used by Departments and those required by the National Accounts.
- 2.47** It would be understandable if those originating the data were less than fully aware of their end use in the National Accounts. People upstream in the chain are often unsure what will happen to the data further down the chain. Equally, people downstream often find it difficult to obtain explanation for apparent oddities and paradoxes in the data they receive. Yet the validity of the data depends on the requirements being communicated to the data suppliers, and on the suppliers documenting the sources and methods of adjustment applied. Of particular importance is the need for contextual information. If, for example, there are large year-on-year changes in data returns, these need to be accompanied by explanatory material. This may involve ONS making changes to the National Accounts to reflect changing methods of service delivery.
- 2.48** In the case of local government expenditure data, the data supplied for local authorities in different countries of the United Kingdom differ in their suitability for purpose. English, Welsh, Scottish and Northern Irish local authorities supply budget and outturn information using different economic categories and with different breakdowns of public services. Consolidation of these data from different sources is dependent on a number of assumptions.
- 2.49** The spending figures need to be deflated to arrive at volume measures. For those functions where output is measured directly, we show in Table 2.4 the deflators employed to arrive at the input volumes, and hence the implied productivity estimates.
- 2.50** Again, the first impression from Table 2.4 is of complexity. The table contains 45 non-blank cells. At the same time, many of the cells employ the same deflator. A small number of aggregate deflators are used repeatedly across different services: for example, the average earnings index for the public sector, and a composite price index of the products purchased by Other Central Government. Moreover, we need to be sure the indices capture in all cases (i) shifts between categories of employee, (ii) additional costs such as National Insurance contributions, (iii) earnings increases in excess of pay settlements, (iv) shifts of activity between functions, and (v) changes in the definition of the status of bodies such as NHS Trusts.
- 2.51** Due to the small expenditure on capital consumption for each function, some specific deflators are unavailable. Where this is the case a generic capital consumption implied deflator for central and local government is used.

Table 2.3 Sources of Data for Output and Input Volume Measures

	Local Government	Central Government
01 General Public Services	<i>Output volume measures are deflated UK expenditure figures for pay, net procurement and capital consumption. Data are supplied by local authorities via ODPM.</i>	<i>Output volume measures are deflated UK expenditure figures for pay, net procurement and capital consumption. Data are supplied via the Treasury.</i>
02 Defence	Not applicable	<i>Output volume measures are deflated UK expenditure figures for pay, net procurement and capital consumption. Data are supplied by MoD via the Treasury.</i>
03 Public Order and Safety	<b>Police – output volume measures are volumes of police activity, crime-related incidents, patrols, traffic incidents etc.</b> Input volume measures are deflated UK expenditure figures for pay, net procurement and capital consumption. Data are supplied by English Local Authorities via ODPM. Coverage is of UK.	<b>Police – output volume measures are volumes of police activity, crime-related incidents, patrols, traffic incidents etc.</b> Input volume measures are deflated UK expenditure figures for pay, net procurement and capital consumption. Data are supplied by Home Office via the Treasury.
	<b>Prisons – not applicable</b>	<b>Prisons – output volume measures are measured directly using total numbers of prisoners.</b> Input volume measures are deflated UK expenditure figures for pay, net procurement and capital consumption. Data are supplied by the Home Office. Coverage is England, Wales and Scotland.
	<b>Probation – not applicable</b>	<b>Probation – output volume measures are measured directly using workload hours of various areas of competence.</b> Input volume measures are deflated UK expenditure figures for pay, net procurement and capital consumption. Data are supplied by Home Office. Coverage is of England and Wales.
	<b>Courts – output volume measures for magistrates courts are measured directly using caseloads of courts weighted average hours or average costs.</b> Data are supplied by Local Authorities via ODPM. Coverage is of England and Wales.	<b>Courts – output volume measures for crown and county courts are measured directly using caseloads of courts weighted average hours or average costs.</b> Input volume measures are deflated UK expenditure figures for pay, net procurement and capital consumption. Data are supplied by Dept of Constitutional Affairs. Coverage is of England and Wales.
	<b>Fire – output volume measures for the fire service are measured directly using number of fires attended and numbers of other services. Coverage is UK, but N.Ireland not included in some indicators.</b> Input volume measures are deflated UK expenditure figures for pay, net procurement and capital consumption. Data are supplied by ODPM.	<b>Fire – output volume measures for the fire service are measured directly using number of fires attended and numbers of other services. Coverage is UK, but N.Ireland not included in some indicators.</b> Input volume measures are deflated UK expenditure figures for pay, net procurement and capital consumption. Data are supplied by ODPM.
04 Economic Affairs	<i>Output volume measures are deflated UK expenditure figures for pay, net procurement and capital consumption. Data are supplied by local authorities via ODPM.</i>	<i>Output volume measures are deflated UK expenditure figures for pay, net procurement and capital consumption. Data are supplied from DTI etc via the Treasury.</i>

Table 2.3 Sources of Data for Output and Input Volume Measures (continued)

	Local Government	Central Government
05 Environmental Protection	<i>Output volume measures are deflated UK expenditure figures for pay, net procurement and capital consumption. Data are supplied by Local Authorities via ODPM.</i>	<i>Output volume measures are deflated UK expenditure figures for pay, net procurement and capital consumption. Data are supplied from DEFRA etc via the Treasury.</i>
06 Housing and Community Amenities	<i>Output volume measures are deflated UK expenditure figures for pay, net procurement and capital consumption. Data are supplied by Local Authorities via ODPM.</i>	<i>Output volume measures are deflated UK expenditure figures for pay, net procurement and capital consumption. Data are supplied by the Treasury.</i>
07 Health	<b>Health – not applicable.</b>	<b>Output volume measures are measured directly using:</b> <b>a) treatment numbers and Reference Costs data from DH. Coverage is England and Wales.</b> <b>b) In addition, further indicator series are used for dental and ophthalmic services with UK coverage.</b> Input volume measures are deflated UK expenditure figures for pay, net procurement and capital consumption.
08 Recreation, Culture and Religion	<i>Output volume measures are deflated UK expenditure figures for pay, net procurement and capital consumption. Data are supplied by Local Authorities via ODPM.</i>	<i>Output volume measures are deflated UK expenditure figures for pay, net procurement and capital consumption. Data are supplied from DCMS etc via the Treasury.</i>
09 Education	<b>Output volume measures are measured directly using pupil numbers in pre-primary, primary and secondary schools obtained from DfES. Also some indicators of numbers of health workers being trained.</b> The coverage is whole UK. Input volume measures are deflated UK expenditure figures for pay, net procurement and capital consumption.	<b>Output volume measures are measured directly using pupil numbers in pre-primary, primary and secondary schools obtained from DfES. Also some indicators of numbers of health workers being trained.</b> The coverage is whole UK. Input volume measures are deflated UK expenditure figures for pay, net procurement and capital consumption.
10 Social Protection	<b>Personal Social Services: Output volume measures are measured directly using: a) numbers of adults in care and home help contact hours obtained from DH b) numbers of children in care from DfES.</b> The coverage is England only. Input volume measures are deflated UK expenditure figures for pay, net procurement and capital consumption.  <b>Social Security: Output volume measures are measured directly for Administration of Social Security using numbers of housing benefit cases.</b> Input volume measures are deflated UK expenditure figures for pay, net procurement and capital consumption. Data are supplied by DWP. The coverage is England, Wales and Scotland.	<b>Social Security: Output volume measures are measured directly for Administration of Social Security using numbers of new benefit claims.</b> Input volume measures are deflated UK expenditure figures for pay, net procurement and capital consumption. Data are supplied by DWP. The coverage is England Wales and Scotland.

DCMS Department for Culture, Media and Sport  
DEFRA Department for Environment, Food and Rural Affairs  
DfES Department for Education and Skills  
DH Department of Health

DTI Department of Trade and Industry  
DWP Department for Work and Pensions  
MoD Ministry of Defence  
ODPM Office of the Deputy Prime Minister

Table 2.4 Deflators that have been used for public service for which volume measures of output are estimated directly

Function/ Service	Pay		Net procurement		Capital Consumption	
	Local Government	Central Government	Local Government	Central Government	Local Government	Central Government
<b>Education</b>	Local government education pay index	Average Earnings Index (AEI) for Public Sector	Implied deflator for total local authority net procurement	Composite price index of the products purchased by Other Central Government	Implied deflator for local authority Education capital consumption	Implied deflator for central government Education capital consumption
<b>Health</b>		The Pay Cost Index (PCI) from the Dept of Health		An aggregate index: 60% of the Labour deflator (PCI) and 40% of the Hospital Services Cost Index (HSCI)		Implied deflator for Health capital consumption
<b>Social Protection – Social Security</b>	Local government pay index	AEI for Public Sector	Composite price index of the products purchased by local authorities	Composite price index of the products purchased by Other Central Government	Implied deflator for local authority Non-Military Administration capital consumption	Implied deflator for central government Non-Military Administration capital consumption
<b>Social Protection – Personal Social Services</b>	Local government pay index		Composite deflator consisting 70% of the Labour deflator and 30% of the procurement deflator (see Health)		Implied deflator for other local authority Health and Social Work capital consumption	

Table 2.4 Deflators that have been used for public service for which volume measures of output are estimated directly (continued)

Function/ Service	Pay		Net procurement		Capital Consumption	
	Local Government	Central Government	Local Government	Central Government	Local Government	Central Government
<b>Public Order and Safety:</b>						
<b>Police</b>	Local government Police pay index	AEI for Public Sector	Composite price index of the products purchased by local authorities	Composite price index of the products purchased by Other Central Government	Implied deflator for local authority Non-Military Administration capital consumption	Implied deflator for central government Non-Military Administration capital consumption
<b>Fire</b>	Local government Fire pay index	AEI for Public Sector	as for Police	as for Police	as for Police	as for Police
<b>Prisons</b>		AEI for Public Sector		as for Police		as for Police
<b>Administration of Justice: Magistrates Courts</b>	Local government pay index		Composite price index of the products purchased by local authorities		Implied deflator for local authority Non-Military Administration capital consumption	
<b>Administration of Justice: Crown and County courts, CPS and Legal Aid</b>		AEI for Public Sector		Composite price index of the products purchased by Other Central Government		Implied deflator for central government Non-Military Administration capital consumption
<b>Administration of Justice: Probation*</b>	(1995-2001) Non-manual administration pay index	(2001-) AEI for Public Sector	(1995-2001) Composite price index of the products purchased by local authorities	(2001-) Composite price index of the products purchased by Other Central Government	(1995-2001) Implied deflator for local authority Non-Military Administration capital consumption	(2001-) Implied deflator for central government Non-Military Administration capital consumption

\* Dates refer to periods of responsibility – probation service responsibility was transferred from local to central government in 2001.

## Conclusions

### 2.52 In conclusion:

- a) Since 1998, ONS has moved progressively towards the replacement of the (output = input) approach by direct measures of the volume of government output. This is an important development. The direct estimates now cover some two-thirds of General Government Final Consumption, which is an impressive achievement.
- b) From the earlier experience in the 1950s and 1960s of the use of the direct measurement approach, we can see that the design of direct output measures needs considerable care and the investment of significant resources. Direct measures of output should be continuously monitored to ensure that they are capturing changes in quality. ONS has to steer a careful course with regard to changes in government policy, guaranteeing the independence of the approach to measuring government output while ensuring that its implementation reflects the realities and circumstances of public spending.
- c) Institutional change in the public sector poses problems for output measurement, and these may be more severe for the direct approach than for the (output = input) convention. Technological change, both specific and general, may not be easily captured, an issue that affects private as well as public services.
- d) The implied measure of productivity for the government sector is obtained from three elements: spending, input price index and direct output measurement. The reliability of all three of these different elements needs to be assessed. The rise in productivity will be understated if the measure of spending is too high, or if the measures of output or price increase are too low, and will be overstated if the reverse conditions hold.
- e) Examination of the recent measures of government output and inputs suggests a number of features in need of explanation.
- f) The process by which the underlying data are assembled is a highly complex one that warrants closer investigation.

# 3 The International Context

---

**3.1** Many users of the National Accounts assume that their form is fully governed by international guidelines. These guidelines, embodied in legal form for EU members by Eurostat (the statistical office of the European Union) are indeed important. The approach taken in these guidelines to the measurement of government output means that, even if ONS had wished to maintain the (output = input) convention, it had no alternative but to consider the development of direct measures. At the same time, the guidelines identify different degrees of conformity, including a distinction between individual and collective public services, and allow some degree of latitude in their implementation. In this chapter, we describe the main features of the guidelines and the experience of other OECD countries. More information is provided in Appendix C.

## **International Guidelines: UN System of National Accounts and OECD Productivity Manual**

**3.2** In this and the next section, we consider in turn the United Nations *System of National Accounts* (SNA), the *OECD Productivity Manual*, the *European System of Accounts* (ESA), and the *Eurostat Handbook on Price and Volume Measures in National Accounts*. An alternative approach, considering the issues in terms of topics (such as definition of the government sector, and the measurement of quality change), is taken in Appendix C. The main features of the different sources of guidance are summarised in Table 3.1.

**3.3** In approaching the guidelines, there are three main questions to which we are seeking answers. The first is the extent to which the international guidelines constrain ONS in its treatment of government output and productivity. As we have already seen, one of the reasons that ONS embarked on the introduction of new measures of government output was the impetus provided by the SNA 1993. The second question is how far the guidelines provide clues as to how we can resolve the issues already identified. Thirdly, the European guidelines form part of an administrative process. To the extent that the United Kingdom has latitude within the legal framework, what are the wider implications of different choices?

Table 3.1 **International Guidance**

Publication	Organisation(s) responsible	Type of guidance on measurement of government output	Status
System of National Accounts SNA (1993), new version SNA 2008 under preparation	UN, OECD, World Bank, IMF, and European Commission. Document prepared by Inter-secretariat Working Group on National Accounts. Approved by UN Statistics Commission	High level guidance	International standard
European System of Accounts ESA (1995)	Eurostat	Fully consistent with SNA 1993, more focused on the circumstances and data needs of the European Union	A legal basis to ensure strict application, providing harmonised statistics
Eurostat Handbook on Price and Volume Measures in National Accounts (2001 edition)	Eurostat	Expansion of ESA 1995 guidance, distinguishing activities, outputs and outcomes. Introduces A/B/C score for methods of member states	Develops ESA 1995 to ensure harmonised price and volume data
OECD Productivity Manual (2001)	OECD	Comprehensive guide to productivity measurement	No formal status, but indicates desirable properties of productivity measures

### UN System of National Accounts (SNA 1993)

- 3.4** The SNA is drawn up by a group on which the major international bodies are represented and is approved by the United Nations Statistics Commission. It is not mandatory, but its recommendations have been taken seriously by all major statistical offices, including Eurostat (see paragraphs 3.14-3.21).
- 3.5** The SNA has an extensive discussion of price and volume measures for non-market goods and services (SNA, 1993, paragraphs 16.133-141). It notes ‘in the case of health and education services provided as social transfers to individual households ... the problems are much less, both conceptually and in practice, than for collective services such as public administration or defence. The objective is to measure the quantities of services actually delivered to households ... For example, individual health services consist of various kinds of consultations and treatments provided to patients’ (SNA, 1993, paragraph 16.135). It goes on to say, ‘The output of the health services needs to be clearly distinguished from the health of the community.’ (SNA, 1993, paragraph 16.136).

- 3.6** For collective services, the SNA recognises that measuring changes in volume of services is distinctly more demanding: ‘it is difficult to measure the output of preventive services, and this is an area in which further research is needed’ (SNA, 1993, paragraph 16.139). It goes on to say, ‘in practice, it may not be feasible to avoid using changes in the volumes of inputs into such services as proxies for changes in volumes of outputs, just as it may sometimes be necessary [to do so] in certain market industries, such as agriculture or construction’ (SNA, 1993, paragraph 16.139). When it is not possible to avoid using an input measure, the SNA states that the input measure should be a comprehensive one, not limited to labour inputs. The SNA then turns to the question of assumed productivity growth where an input measure is employed. ‘A possible alternative method [makes] an explicit assumption about changes in labour productivity: for example, that labour productivity grows at one per cent per year in the production of the non-market service in question. An assumption of zero productivity growth is the most common one in practice because it is felt to be more neutral, even though it is inevitably somewhat arbitrary’ (SNA, 1993, paragraph 16.141).
- 3.7** The position taken in the SNA 1993 was cited by ONS (*Economic Trends*, February 1998) as one of the main motivations for the programme of work on direct measures; and the approach adopted seems in line with that envisaged in the SNA at that time.

#### SNA revision 2008

- 3.8** The process of revising the SNA 1993 is already underway. The work programme is being steered by the Advisory Expert Group (AEG) on National Accounts of the Intersecretariat Working Group on National Accounts. The AEG, in which ONS participates, met at the International Monetary Fund in February 2004, and is scheduled to have four further meetings, with the aim of adopting the proposed changes in February 2007.
- 3.9** Of direct relevance to this review is an item on the tentative agenda for November 2004: ‘government owned assets – cost of capital services’. As noted below, the OECD Productivity Manual indicates that the volume of capital services is the appropriate measure of capital input for productivity analysis, and it has a solid theoretical basis in the work of Dale Jorgenson (1963 and 1995), and others. So far, relatively few countries have made estimates of the flow of capital services (Schreyer, 2003), but progress is beginning to be made. The Australian Bureau of Statistics (ABS) publishes two distinct and complementary capital measures, and stands out in that it ensures full consistency between the different measures: (1) a measure of capital services, as part of ABS multifactor productivity series; and (2) an end-year net capital stock, as part of the Australian System of National Accounts.
- 3.10** There appear to be long lead times in introducing new issues into the revision process for the SNA. We quite appreciate the problems in securing agreement among a large number of countries; we also understand that implementation is itself glacial in its pace. Nonetheless, the issues addressed in this review appear pertinent to the revision of the SNA, and we hope that the ideas put forward here will be allowed to enter the process.

### OECD Productivity Manual

- 3.11** The OECD Productivity Manual was produced with the objective of providing an accessible guide to productivity measurement, particularly for use in statistical offices. It identifies desirable characteristics of productivity measures and, although there is no strongly prescriptive element, it seeks to improve international harmonisation. Strong links are made to the underlying economic theory.
- 3.12** The Manual covers the whole of the economy, but it has lessons for our specific concern with the government sector. One important point is that the term ‘productivity’ is used in different ways, and we need to be precise in its application. (This has been clearly recognised by ONS in its work: see for example, *Economic Trends*, May 2002, pp 21.) When people talk about the input approach to the public sector ‘assuming zero productivity growth’, this could have several interpretations; and correspondingly the (output = input) approach can have several interpretations. For example, in some cases, the inputs used to estimate output are confined to employment. In this case, it is output per unit of labour input (‘labour productivity’) that is being assumed constant. In other cases, it is total inputs that are used, in which case it is total input productivity that is being assumed constant. Labour productivity in the public sector may be rising because more capital per worker is being applied. (Maurice (1968) refers to the introduction of computers.) If public sector output was based only on labour input, then the increased output would be missed; if public sector output was based on total factor input, then the increased output might be captured.
- 3.13** The manual is the acknowledged source of methodological guidance for estimating inputs in volume terms. It recommends that:
- for labour inputs, the volume measure should be hours worked by workers at different skill levels, weighted together using expenditure on each of these categories;
  - for intermediate consumption, purchases of goods and services of different types should be deflated using an appropriate price index for each category, i.e. an index constructed to reflect changes in the quality of the goods or services measured; and
  - for capital consumption, there should be a method such as a perpetual inventory model that estimates it.

### European System of Accounts (ESA 95) and the Eurostat Handbook on Price and Volume Measures in National Accounts

- 3.14** ‘The economic accounts in real terms, i.e. adjusted for price changes, are a fundamental tool for analysing a country’s economic and budgetary situation, provided they are compiled on the basis of unique principles that are not open to different interpretations.’ (European Commission Decision of 17 December 2002).
- 3.15** The ESA 95 was adopted in the form of a Council Regulation Number 2223/96, dated 25 June 1996. There have been subsequent amendments, and particularly important here has been the publication by Eurostat in 2001 of the *Handbook on Price and Volume Measures in National Accounts*, which seeks to provide a complete discussion of the issues involved in measuring current and constant price quantities, from general principles to the deflation of individual goods and services. The recommendations were embodied in the European Commission Decision of 17 December 2002. Implementation is due for the accounts covering 2006 data, although Denmark has secured a derogation until 2012. A primary objective of this Commission Decision is to harmonise measures of GDP growth, which is a key concern of this review.
- 3.16** Central to the approach of the Eurostat Handbook is the introduction of an A/B/C classification. This distinguishes between:
- A methods: the most appropriate;
  - B methods: methods which can be used where it is not possible to apply an A method; and
  - C methods: methods which should not be used.
- 3.17** In particular, ‘C methods are too far away from the ideal to be acceptable. They would generate too great a bias or would simply measure the wrong thing’ (Eurostat Handbook, section 1.4). The Commission Decision means that from 2006 (or, in some cases, 2004 or 2005), C methods are no longer allowed.
- 3.18** This classification applies to the whole of the National Accounts, not just to the government sector, but the Eurostat Handbook discusses its application to non-market output, covering both individual services (those consumed by individual households) and collective services provided to the society as a whole. The Eurostat Handbook rejects the convention that (output = input) on the grounds that it ignores all changes in productivity. It rejects, too, the input approach complemented by an assumed productivity change: ‘there is no reason at all why for example a one per cent productivity adjustment would be more plausible than a zero per cent adjustment. Productivity might just as well have declined’ (Eurostat Handbook, section 3.1.2.1). The Eurostat Handbook equally addresses the argument that the input method, if not ideal, can at least ensure comparability across member states if all make the same assumption about productivity. It is robust in dismissing this argument: ‘a harmonised assumption about productivity does nothing to make the resulting estimates of output more comparable. The more different the developments in productivity among member states, the less comparable are the results from using the same productivity change assumption’ (Eurostat Handbook, section 3.1.2.1).

- 3.19** Member states are therefore required to develop direct output measures. The requirements for different gradings are summarised in Table 3.2. The criteria for an A grade provide a valuable checklist when considering the individual functions in ONS accounts. The B classification still represents compliance. In the case of collective services, the B classification does allow input methods, provided that they estimate the volume of each indicator separately, taking quality changes of inputs into account. Provided that this criterion is satisfied, it would be open to ONS to continue to use input measures for collective services, although it could not achieve an A classification. For individual services, input measures are unacceptable. It should be noted that the classification of services into collective and individual is an issue. The Handbook gives examples but some are open to question.
- 3.20** In considering alternatives to inputs, the Eurostat Handbook distinguishes between activities, outputs, and outcomes. These are described in more detail below:
- a) *Activity* is, ‘for example, the number of operations in hospitals or number of patrols carried out by the police. Such data can often be found. Activity indicators reflect what the non-market units are actually doing with their inputs and are therefore closer to the output. However, suppose for example that new improved forms of medical treatments reduce the number of operations necessary. Taking the number of operations, as an indicator would imply a decrease of output and productivity, which does not seem appropriate in this case. Using activity indicators often does not lead to reasonable productivity numbers. However, for some collective services, activity indicators may be the only indicators that can be found’ (Eurostat Handbook, section 3.1.2.1).
  - b) *Output* ‘is the preferred approach. However, it is not always easy to define exactly what the unit of output is. For individual goods and services it is in principle possible to define the output, since an actual delivery of that output takes place from the producer to the consumer(s) ... For example, for education, the output is the amount of teaching consumed by a pupil. For hospital services, the output is the amount of care received by a patient. For cultural services, the output is the amount of theatre plays consumed. For collective services, however, there is no transaction between producer and consumer since these are provided simultaneously to the society as a whole. It becomes therefore very difficult to define the output. It is very difficult to say for example what the unit of output is of defence or police services.’ (Eurostat Handbook, section 3.1.2.1).
  - c) *Outcomes* are ‘for example indicators of the level of education of the population, life expectancy, or level of crime. Such indicators might be influenced by factors that are unrelated to the activity, and therefore are generally not representative of the output. In some cases, however, outcome indicators can be used as indicators for the quality of the output’ (Eurostat Handbook, section 3.1.2.1).
- 3.21** These are valuable distinctions, and they have informed the approach taken by ONS (see Chapters 7 to 10 below for discussion of their application to individual spending areas). At the same time, the distinctions need some elaboration, and this will be our starting point in Chapter 5.

Table 3.2 Eurostat Handbook Recommendations for Government Output

Type of service	A/B/C methods
Individual services (such as Education, Health, Social security, Recreation and Culture)	<p>A methods – output indicator approach where the indicators satisfy the following criteria:</p> <ul style="list-style-type: none"> <li>a) They should cover all services provided;</li> <li>b) They should be weighted by the cost of each type of output in the base year;</li> <li>c) They should be as detailed as possible;</li> <li>d) They should be quality adjusted.</li> </ul> <p>B methods – output indicator approach where the criteria are not fully satisfied: eg the level of detail could be improved or the measure does not take into account changes in quality.</p> <p>C methods – if input, activity or outcome is used (unless outcome can be interpreted as quality-adjusted output) or if coverage of output method is not representative.</p>
Collective services (such as General Public Administration, Defence, Police services, and Research and Development).	<p>Broadly the same as for individual services but:</p> <p>B methods – input methods are B methods, as are the use of volume indicators of activity. If input methods are used they should estimate the volume of each indicator separately, taking quality changes of inputs into account. Applying productivity or quality adjustments to the sum of the volume of inputs is not recommended.</p> <p>C methods – the use of a single input volume indicator is not a B method.</p>

### Progress of Other Countries in Measuring Government Output

- 3.22** In 2003, ONS carried out a simple fact-finding survey of the steps being taken in other countries to implement the SNA 1993 recommendation regarding the use of direct volume measures. Questionnaires were sent to many OECD countries, and form one input into this section of the report. The second input has been the consultations undertaken with the statistics offices of Australia, Finland, Italy, the Netherlands, Norway, Sweden, and with Eurostat and OECD. (See Appendix D.)
- 3.23** The United Kingdom is among the world leaders, both in terms of its extensive use of direct volume measures of output, covering two-thirds of government output, and in the priority attached to developing this dimension of the National Accounts. Australia, Italy and the Netherlands also place a priority on developing direct volume measures, and they cover between 20 per cent and 50 per cent of government output. New Zealand covers 60-70 per cent of central government output.

- 3.24** Several countries, including Finland, Germany and Norway, plan to incorporate them for certain sectors in their National Accounts in the next few years, although it should be noted that Germany believed that the traditional input measures were more suitable generally, their view being that they offered international comparability and timely data. The experience of the Netherlands suggests that GDP growth is lower using direct volume measures; Italy has seen no systematic impact; Australia, New Zealand and Canada found that estimates of GDP growth rates increased with the introduction of direct volume measures.
- 3.25** One of the key issues is how to capture quality changes in the government output. In the case of Education, a number of countries make adjustments, but little account is otherwise taken of quality in the measures actually employed. Several countries reported that they are seeking to develop methods to take account of quality change.
- 3.26** The United States does not make output-based estimates of government output, despite the fact that it has had extensive experience of seeking to measure government productivity. From 1973 to 1994, there was a federal productivity measurement programme, covering two-thirds of federal civilian employees, and tracking a sample of state and local government activities, with data going back to 1967 (Fisk and Forte, 1997). The programme built up some 2,500 indicators of productivity, covering such items as medical care provided (weighted composite), letters delivered, acres of fine lawn maintained, cases disposed by the courts, and disaster loans approved. There has been substantial academic research on productivity change: see, for example, Hulten (1984). But the productivity measurement series have not been used in the construction of national accounts. In part, this reflects the fact that the federal productivity measurement programme was the responsibility of the Bureau of Labor Statistics, whereas responsibility for the US national accounts is with the Bureau of Economic Analysis at the Department of Commerce. In part, it reflects the fact that their objectives were rather different. But there have been voices calling for a different approach. In his magisterial survey of extended national accounts, Robert Eisner criticised the US official treatment of government services as follows: 'Government output is counted on the basis of market inputs of labor services ... The value of government output is understated by the extent to which it ignores inputs of capital and land. And there is no imputation for the value of labor time not paid for.' (Eisner, 1988, pp 1,620).

## Summary

### 3.27 To summarise:

- a) The programme of work initiated by ONS in 1998 was in harmony with the SNA 1993 guidelines; other countries are engaged in the process of revising their methods for estimating government output, although the United Kingdom leads the way in terms of the extent of use of direct volume measures.
- b) It is hoped that the work of ONS on the output of the government sector will influence the SNA 2008.
- c) The ESA 1995, and the *Eurostat Handbook on Prices and Volumes Measures in National Accounts*, have expanded the SNA guidance, introducing an A/B/C classification. By 2006, C methods will no longer be acceptable under a European Commission Decision of 2002. In the case of individual services, this precludes use of the (output = input) convention in measuring government output; in the case of collective services, input measures may be retained as a B method providing that they satisfy certain criteria.
- d) In a number of countries, there are concerns about the adequacy of the current direct output measures, and in particular the treatment of quality change.

# 4 The Functions and Limits of National Accounts

---

- 4.1** *'The System of National Accounts (SNA) consists of a coherent, consistent set of macroeconomic accounts and tables designed for a variety of analytical and policy purposes.'* (SNA, 1993, paragraph 1.68). *'There are two aspects that need to be separated. The first is the adequacy of the main aggregates of the System as summary indicators of economic activities taking place within the economy as a whole ... The second is the more general question of the validity of using measures of aggregate production or consumption as indicators of welfare'* (SNA, 1993, paragraph 1.70).
- 4.2** The authors of the SNA correctly emphasise the variety of purposes served by measures of national income. The introduction of official national accounts in the United Kingdom during the Second World War may indeed be seen as emanating from two related, but different, streams of economic thought. The first, and the most pressing at that time in terms of policy needs, was the development of economic management at the macroeconomic level. It was no accident that Keynes became a strong advocate of the national accounts being developed by Meade and Stone. The second is the expression of the level of national welfare in terms of national income, stemming from the earlier, welfare economic tradition symbolised by Pigou, developed by Hicks, Samuelson and others, and implemented by Bowley and Clark.
- 4.3** In what follows, we consider in turn these two influential strands, considering how changes in the method of estimating government output would affect the uses of national accounts. We then go on to examine the relation with other possible uses.

## Macroeconomic Management

- 4.4** National accounts provide data on the major economic flows necessary for aggregate economic management. As it was put by Christopher Allsopp in his *Review of Statistics for Economic Policymaking*, 'the demands of monetary and fiscal policy establish a clear need for high-frequency, macroeconomic data' (Allsopp, 2004, paragraph 2.9). These data are used to estimate causal relationships at the macroeconomic level, and they furnish the input into macroeconomic models used to make forecasts and aid the formation of economic policy.
- 4.5** In simplified terms, the central issue of macroeconomic management may be seen in terms of balancing aggregate supply with aggregate demand. Considerable attention is paid by the Monetary Policy Committee of the Bank of England (MPC) and by the Treasury to the 'output gap' between potential supply and demand in the economy. Our particular concern here is with how the measurement of government output affects this gap.

- 4.6** In presenting the February 2004 *Inflation Report*, the Governor of the Bank of England, Mervyn King, said ‘in assessing inflationary pressures, the official GDP data may not be the best guide to the balance between demand and supply in the economy as a whole. GDP includes an estimate of the output of the public sector. That is extremely difficult to measure in sectors such as health and education’. We fully share the view that there is no one single number that will serve all purposes, and that different aggregates are relevant to answering different questions. As it was put by Sir John Hicks, ‘there may be more than one *Money Value of the Social Income*, each corresponding to a different purpose of calculation’ (Hicks, 1940, pp 106).
- 4.7** From the standpoint of macroeconomic policy, it may therefore be that the appropriate aggregate measures are those for the private economy, excluding government from both the supply and demand side of the equation. The main concern is then to make an appropriate subtraction from potential supply of that part of resources absorbed by the public sector. To continue the quotation from Mr King, ‘what matters is not the value of the services provided by the public sector but the opportunity cost of the resources that would otherwise be employed in the private sector’. This ‘resource cost’ measure can be estimated by adding the public sector’s procurement to a hypothetical quantity of goods and services that would have been produced by the public sector workforce if it had been employed in the private sector (as calculated by the Bank of England in the May 2004 *Inflation Report*).
- 4.8** In terms of the three ingredients identified in paragraph 2.23, what matter are the public expenditure on inputs (and the appropriate price indices), but not the direct measure of output of the government sector. If we were to discover that the measure of spending is too high, or the price index is too low, then the government would be making less demand on resources than at present indicated by the National Accounts. The converse would be true if we were to discover that the measure of spending is too low, or the price index is too high. The estimate of spare capacity by the MPC would then be affected.
- 4.9** If government outputs are separable in this way from private supply and demand, then this allows considerable simplification. Any recommendations that we make concerning the direct measurement of government output should not therefore affect the macroeconomic policy stance. At the same time, if our investigation suggests that changes need to be made to the construction of the nominal spending figures or of the associated input price indicators, or both, then this will have implications for macroeconomic management.
- 4.10** One evident conclusion is that users of the National Accounts need to be aware of the different elements that in principle (not necessarily in detail) enter the different macroeconomic variables.

- 4.11** The above argument assumes, in effect, a degree of separability between public output and the private economy. To a first approximation, this seems a reasonable assumption in terms of short-run macro-management. Taking a longer-term view, there may be feedback effects. If, for example, smaller class sizes are leading to better-educated school children, then there will eventually be a better-educated labour force, representing an increase in effective labour supply. Improved medical care may have a more immediate impact through reduced sickness absence.

### Assessing Overall Economic Performance and Welfare

- 4.12** ‘Certain key aggregates of the System, such as GDP and GDP per head of population, ... are widely used by analysts, politicians, the press, the business community and the public at large as summary, global indicators of economic activity and welfare’ (SNA 1993, paragraph 1.68).
- 4.13** The second primary use is as an indicator of the contribution of economic activity to increasing welfare. In this respect, both public and private economic activities play their role. The contribution of private goods and services is perhaps more evident. Market prices ‘represent the relative value to the individual of different goods and services, on the usual assumption that the price paid for each commodity is proportional to its marginal significance. Thus an increase in the domestic product at market prices, other things being equal, can be regarded as indicating *prime facie* an increase in the ‘economic welfare’ of the community.’ (Maurice, 1968, pp 14). The SNA 1993 makes the same point under the heading ‘Changes in welfare’ (SNA, 1993, paragraph 1.76), the title of which reminds us of the large welfare economic literature of the 1940s and 1950s (for example, Samuelson, 1950) on the conditions under which we can identify an increase in real income.
- 4.14** A substantial part of public output takes the form of services provided to individuals: for example, a GP consultation. There may be no market price, but the service adds to individual welfare in the same way as a consultation with a vet, which is paid for as a market service, adds to the pet owner’s satisfaction. Other public sector output takes the form of a collective good, where benefits are consumed jointly, and non-excludably, by the whole population. In the case of a pure public good, there is again a *marginal* value to each consumer; and, as was explained by Samuelson in his pure theory of public expenditure (Samuelson, 1954), we have to add these marginal valuations to arrive at the total value of the activity. (With public goods, we add demand curves vertically for a given quantity; whereas with private goods we add the demands at a given price.) It should be noted that we are concerned in both cases with the marginal benefit, not with the total benefit. In the case of private goods, we are taking the price as an indicator of the willingness of the consumer to pay for the marginal unit; we are not taking into account the consumer surplus associated with intra-marginal units. National income does not measure the area under the demand curve but the rectangle formed by the price times the quantity. In the same way, with public goods, we are adding the marginal valuations not the total consumer surplus.

- 4.15** In considering the welfare interpretation, there are two possible misunderstandings. The first is that national income provides a total measure of welfare. This is not the case, as is clearly stated in the SNA (1993), ‘the consumption of goods and services, both individually and collectively, is one of the most important factors influencing the welfare of a community, but it is only one of several factors. There are also others, such as epidemics, natural disasters or wars, that can have major negative effects on welfare, while others, such as scientific discoveries, inventions or simply good weather, may have significant positive impacts. These factors obviously do not enter into the measurement of GDP, which refers only to the flow of goods and services produced within a given period.’ (SNA (1993), paragraph 1.69). The calculation is therefore different from extended measures of ‘net economic welfare’ devised by James Tobin and others. National income is an indicator of the *contribution* to welfare of a specified set of economic activities. In the next chapter, we make this more concrete.
- 4.16** Secondly, welfare is only one of the considerations entering into policy judgements. Governments have multiple goals. They are concerned with distribution as well as with totals. They are concerned with rights and procedural justice, as well as with outcomes. Equity also has a value for governments, but it is not captured in the National Accounts and nor should it be. For these reasons alone, there is no reason to expect government policy to be directed solely at maximising national output; nor, conversely, should the output measure be determined solely by the policy objectives.

### Relation with Government Performance Targets

- 4.17** When ONS embarked on its new approach to the measurement of government output, it made quite clear that there is a difference between National Accounts estimates of output, on the one hand, and performance measures for the management of public services on the other hand. Neuburger and Caplan (1998) recognised that performance measures and output indicators will use much of the same data, but spelled out the differences in the requirements. In the case of performance indicators, they argued, the need was for precise, transparent and simple measures, not subject to manipulation, but there was no requirement for stability over time, and they could be selective in their coverage. In contrast, the output measures for National Accounts purposes need to be as comprehensive as possible and to be consistent over time. Moreover, while simplicity and transparency of compilation would be desirable, national accountants typically expect to have to make complex adjustments to raw data.

- 4.18** The difference between national income measurement and performance measures is related to the distinction between outputs and outcomes. Examination of the Public Service Agreement (PSA) targets of UK government departments shows that a number are concerned with total outcomes. (Our quotations are all from 2002 PSAs.) In the case of the Department of Health, for instance, PSA target 6 is to reduce substantially the mortality rates from the major killer diseases. This is a total outcome target. The same applies to PSA target 9 (reducing the under-18 conception rate) and PSA target 11 (reducing health inequalities). As we have already noted, national income measures the contribution of activities to outcomes, and these may be only part of the story. It could be, for instance, that rising affluence is leading to over-eating and that the government's health promotion activities are swimming against a powerful tide. We would not want to measure its output as negative because health status is worsening; rather we wish to measure how far it has succeeded in offsetting the negative effect of other factors.
- 4.19** The two activities are different. National Accounts provide indicators of broad trends; to try to use them as microeconomic measures of public sector performance misunderstands their nature and limitations. The National Accounts indicators need to be coherent with the evidence from performance studies. But they are not a substitute. The construction of National Accounts may, as noted above, make use of similar data sources. However, there are risks, clearly recognised by ONS, that 'the use of data for a performance indicator may transform its meaning as it becomes a target' (Neuburger and Caplan, 1998, pp 35). Circumstances where National Accounts measures would provide perverse incentives in respect of the good management of the public resources concerned need, at the least, close scrutiny.

### International Comparisons and Obligations

- 4.20** From the early days of National Accounts, there has been considerable interest in international comparisons, and this was one of the major reasons for the United Nations to initiate the *System of National Accounts* more than half a century ago. The data are used to compare economic performance across countries and league tables of national income often take on a political significance. As was noted earlier, different approaches to measuring government output may lead to different measures of growth rates, these differences being purely statistical.
- 4.21** More recently, the United Kingdom's interest in such comparisons has taken a particularly concrete form. With the development of the European Union, national accounts data have come to play a legal role in the determination of budgetary payments, leading to the establishment of the ESA guidelines with the force of Commission Decisions. National accounts data play a key role in EU macroeconomic policy. As it was put in the Foreword to *European System of Accounts*, 'to achieve the objectives set by the Treaty on European Union, and more specifically Economic and Monetary Union, we need high-quality statistical instruments which provide the Community institutions, governments and economic and social operators with a set of harmonised and reliable statistics on which to base their decisions' (ESA, 1996, pp I).

**4.22** These considerations mean that we have to consider the implications of EU membership for any changes in the measurement of government output. The short answer is that this does not appear a matter for concern. UK contributions to the European Union are determined by current price Gross National Income figures. They would be affected by any reclassification of final government expenditures as intermediate expenditures, but not by changes to the constant price GDP series under consideration here.

### Micro-Macro Links

- 4.23** The SNA 1993 refers to ‘micro-macro links’, concluding that, while definitions and concepts should be as far as possible consistent between macroeconomic accounts and micro-databases, they cannot in general be fully integrated. There may, however, be some degree of ‘read across’ between macro- and micro-analyses.
- 4.24** In the present context, we should refer to the studies, conducted by ONS (CSO) since the 1960s on ‘the effects of taxes and benefits on household incomes’, published annually in *Economic Trends*. Using data on individual households, drawn from the Family Expenditure Survey, ONS estimates the distributional impact of direct and indirect taxes, cash transfers, and benefits in kind provided free by the government. It is the last of these that is relevant here. Among the benefits in kind covered by the study are state education, school meals and welfare milk, and the National Health Service.
- 4.25** Education benefit in the taxes and benefits study is estimated from information provided by the Department for Education and Skills on the cost per pupil in primary, secondary, special schools, universities, and further education establishments. The value of the benefit attributed to the household depends on the number in the household recorded as receiving each kind of state education. This effectively adopts an (output = input) approach.
- 4.26** In the case of health care, the study uses the average cost of different types: hospital inpatient/outpatient care, GP consultations, dental services, etc. Each individual in the survey is then allocated a benefit from the NHS according to the estimated average use made of the various types of health service by people of the same age and sex. The important feature is that this is an ex ante allocation, rather than an ex post measure of the benefit actually derived. We return to this distinction in the next chapter.
- 4.27** The taxes and benefits studies have been criticised by a number of authors, such as Le Grand (1982). A number of these criticisms concern the distributional impact: for example that there is differential use of the health service by different income groups. Other criticisms apply to the total allocated, notably that the benefits of education are not fully captured by the cost of the inputs.

## The Limits of National Accounts: Conventions and Margins of Error

- 4.28** The National Accounts have become so much part of the economic life of the country that their limits are often overlooked.
- 4.29** First, we should stress that the definition of national income is a matter of *agreed conventions*. This will have been obvious from the account given in the previous chapter of the international guidelines. As described in the CSO *National Income Statistics: Sources and Methods*, ‘the meaning to be attached to the aggregate of national income, product or expenditure is essentially arbitrary and limited. The comprehensiveness of the aggregates is limited by convenience and convention; the valuations placed on goods and services ... do not provide precise measures of changes either in welfare or in productivity.’ (Maurice, 1968, pp 15). This does not mean that they are worthless. The quotation continues to say that ‘the significance of the broad trends shown by the aggregates is often unmistakable.’ (Maurice, 1968, pp 15).
- 4.30** This means that the findings of the National Accounts have to be interpreted in the light of these conventions. Particular care is needed at a time of change, as is well illustrated by the measurement of government output. At the present, the United Kingdom has introduced direct output measurement for the government sector for a wide range of spending functions; other countries have introduced it less extensively or not at all. This means that measures of real growth rates are not comparable. Even after 2006 when compliance is due for all member states (apart from Denmark), there will still be countries using B methods rather than A methods.
- 4.31** Secondly, national income is an *indicator*, not a precise addition over all the possible constituent parts. The origin of early academic estimates of national income (such as those of Bowley) in the tables of incomes assembled by the Inland Revenue may lead people to suppose that ONS is actually measuring the total of national income in the same way as individuals measure what has flowed into their bank accounts. Even at that time, with a narrow definition of personal income, such aggregates were estimates, extrapolating from a limited base of data.
- 4.32** This brings us to a third point: that the National Accounts variables are *measured with error*. When ONS publishes a figure stating that the government output at 1995 prices was £156.4bn in 2001, this is an estimate surrounded by a considerable margin of error. This should be obvious, but the fact is often overlooked and needs to be stated explicitly. Moreover, national accounts are a system, but different components are measured with differing degrees of accuracy. In order to complete the system, elements have to be included which are less reliable than others. The reader may well ask whether the figure just cited for government output is as reliable as the private sector element of constant price GDP.

**4.33** The significance of errors of measurement depends on the purpose for which the accounts are to be used. In the present context, focus on the growth of constant price output means that an unknown but fixed bias in the estimate would not matter. If the input into the direct output measure were to be always understated by the same percentage then we do not need to know the percentage, since we are using only the changes over time. (This is another way of expressing the point that the direct output measures provide no information on the *level* of output.)

### Conclusions

**4.34** Our conclusions are:

- a) The National Accounts serve several purposes, and no one single number will serve all purposes; different aggregates are relevant to answering different questions; what is essential is that the relation between these aggregates be clearly presented.
- b) National income is an indicator of the contribution to welfare of specified economic activities; it is not a measure of total economic welfare; aggregate welfare is not the only objective of government policy.
- c) The National Accounts estimates for the government sector do not have the same purpose as microeconomic measures of public sector performance; to try to use them in the same way as public sector performance targets misunderstands their nature and limitations.
- d) The National Accounts are built on a series of agreed conventions; they are subject to margins of error that vary across different parts of the National Accounts; the significance of these errors depends on the purpose for which the figures are used.

# 5 A Methodological Framework

- 5.1** This chapter sets out a framework for considering the principles to be applied in developing the present measures of the change over time in government output at constant prices. The international guidelines described in Chapter 3 provide a point of departure, but a great deal remains to be done before the intentions signalled in these documents can be fully realised. The United Kingdom is in the vanguard of countries seeking to implement the European Commission Decision, and in some cases we are breaking new ground. In view of this, we feel that it would be valuable to enunciate a set of principles on which to base the measurement of the output from government spending. These principles build on the current practice, often serving only to make explicit the implicit methodological approach, but a clear statement may inform the professional and public debate. The main principles we recommend are set out below, and cover the direct measurement of output, the measurement of inputs, and the measurement of productivity.
- 5.2 Interim Conclusion (1) Regarding Principles:** The direct measurement of the output from government spending, and the measurement of inputs and productivity, should be based on a set of principles, within the framework set by international guidelines.

## The Concept of Added Value

- 5.3** The input/activity/output/outcome distinctions made in the Eurostat Handbook seem an excellent starting point. To remind the reader, and taking the health service as an example, we identify the *inputs* as the time of medical and non-medical staff, the drugs, electricity and other inputs purchased, and the capital services from the equipment and buildings used. These resources are used in primary care and hospital *activities*, such as a GP making an examination or the carrying out of a heart operation. These activities are intended to embody health care designed to benefit the individual patient. To the extent that they do, the health care provided constitutes the *output* associated with these input activities. Finally, there is the health *outcome*, which may depend on a number of factors apart from the output of health care, such as whether or not the person gives up smoking.
- 5.4** Inputs are not an appropriate measure, for reasons already made clear, and, while activities may be the only available indicator and hence have to be used, they, too, are an intermediate variable. The choice between output and outcome, on the other hand, is less obvious, and indeed the definitions appear less than fully transparent. In the case of Education, for instance, ‘the output is the amount of teaching consumed by a pupil’ and an outcome is ‘the level of the education of the population’. But one could argue that the amount of teaching is a measure of activity, not of output, and that the real output is the increase in the level of education attributable to that teaching. The Eurostat Handbook later refers (Eurostat Handbook, section 4.12) to output as ‘the quantity of

teaching (that is, the transfer of knowledge, successfully or not)', which again sounds more like an activity. To take an analogy from marketed goods, this seems parallel to the number of bricks put into the kiln. If the bricks come out properly fired, then the purchaser is willing to pay for them, and this is recorded as a market output. But the customer is not willing to pay for those that come out cracked. A restaurant may offer a splendid menu, but if no customers place an order, then there is no marketed output. It does not seem reasonable to treat time wasted in the classroom as an output, other than as serving a purely custodial function. The Eurostat Handbook describes 'the skill and knowledge that a pupil achieves' as an outcome, but the contribution of schools to this skill and knowledge looks a closer counterpart to marketed output.

- 5.5** Approaching the question another way, we may note that the objection to the use of outcome is that the status of the population is affected by many factors other than public spending. Parents devoting more time to teaching their children increase the level of education of a society in a way that cannot be attributed to public spending. This objection is well based. But it does not rule out the approach above. What it does suggest is that what we want to measure is the *incremental impact on outcomes arising from the activities of the public sector*. In the case of Education, the objective should be to measure the improvement in individual educational outcomes attributable to the schools.
- 5.6** This approach underlines the parallel with what the National Accounts seek to measure in the case of market output. In public sector terms, it allows us to bring outcome into the picture, while emphasising that it is the incremental contribution and not the level that is being measured. This helps ensure coherence with the departmental performance indicators, since it is outcomes that much concern departments. At the same time, we hope that this preserves the spirit of the international guidelines. We *are* looking at the educational level of the population but in terms of how it is affected by the activities of the education sector. Added value is increasingly used in studies of school performance. It helps shift attention from physical quantities (pupils in school) to the notion that the value depends on quality and context (the enhancement of future earning power).
- 5.7** Therefore, **Principle A**: the objective should be to measure output from government spending in terms of its incremental contribution to individual or collective welfare, in the same way as market output.

### Parallel with the Private Sector

- 5.8** The thrust of the SNA 1993 was, as emphasised by Neuburger and Caplan (1998), 'to treat, as far as possible, public output in the same way as private output: the same general procedure can be used in both the public and the private sector' (*Economic Trends*, 1998, pp 31). This seems clearly right. The issues of measuring output and productivity apply across the National Accounts as a whole, and the principles applied to their measurement should as far as feasible be the same. This is particularly important in view of the transfers of activity that have taken place across the private/public boundary. It is evidently desirable that the relocation of an activity does not in itself lead to a change in the estimate of national output.

- 5.9** Once we consider the public sector as supplying services, either individual or collective, it is therefore reasonable to ask how far we can simply borrow from private sector experience, with appropriate modification? Our terms of reference identify the need for comparability with measures of private sector services output and costs. We have indeed already found the analogy useful in the preceding section.
- 5.10** In the private sector, the concept of net output with which we are concerned is the value added: i.e. gross output less the goods and services bought in from other consumers. In practice, it is often difficult to measure the value added directly, and proxy indicators are used for the changes in value added. One such approach is to use the gross turnover, deflated using a relevant price index, on the assumption that gross output and value added move in parallel. Such an assumption would not be warranted if there were major changes in the industry: for example, a move to greater out-sourcing.
- 5.11** The proxy method just described has no counterpart in the case of the public sector, since there are no sales, but two other methods do have relevance. The first is to use the number of employees in employment to represent changes in output. This is the method currently used for private educational institutions, whose constant price value added is based on the number of full-time equivalent employees. According to the CSO in 1985, where employment is used in the private sector, ‘arbitrary allowances for changes in labour productivity in these private industries have been made for 1973 onwards (with effect from the 1984 *Blue Book*). This procedure is thought to give better proxies for output than the previous implicit assumption of no change in labour productivity.’ (CSO, 1985, pp 39). This method is, however, the (output = input) approach that we are seeking to leave behind.
- 5.12** The second method is to take direct output measures. For example, in the case of post and courier services, there are 14 indicators, such as the number of first class letters, tonnes of overseas letters, use of courier services, which are combined with appropriate weights. This method is similar to that already employed by ONS in the direct measurement of government output, and it seems to have some of the same problems. According to the Eurostat Handbook, ‘volume indicator methods based on detailed indicators of the many types of services provided, for example number of letters/parcels broken down by different postage rate, are B methods’ (Eurostat Handbook, section 4.8.3.1).
- 5.13** To summarise, while it is important to ensure consistency with private sector practice, so that national output is not materially affected by changes in the location of activity, the estimation of private sector service output appears to be grappling with a number of the same problems.

## The Coverage of Output Indicators

- 5.14** In considering coverage, it is important to distinguish between the coverage of different government functions and the coverage within functions. In the current UK context, the first concerns the third of total government spending not yet covered by direct output measures; the second concerns the coverage within functions such as Education. We begin with the coverage *within functions*.

### Adequacy of coverage within functions

- 5.15** The fact that national income is an indicator, not a precise addition over all the possible constituent parts, means that we should not expect an indicator of government output and inputs to be based on data that cover every single pound of government spending under a particular heading. At the same time, the international guidelines are right to stress the need for extensive coverage. ONS practice has shown, for example in the new Health measures, that a high level of coverage can be achieved. The Eurostat Handbook requires for the A grading that the indicator should cover all services provided. That is a counsel of perfection, and realistically a balance has to be drawn between cost and coverage. What we would expect is that, within a government function, all major services should be covered.
- 5.16** To summarise in **Principle B**, the procedure of defining direct output indicators within a government function should start by seeking to identify the services provided by government to households and firms, and attempts made to find data to reflect these services as comprehensively as possible (rather than working back from available indicators). Where, initially, it is necessary to apply an indicator from another service, this should be explicit. The coverage of indicators within a function should be reassessed on a regular basis.

### Extension of coverage by function

- 5.17** The extension of coverage by function needs to be considered separately for individual and collective services. We have already seen that the United Kingdom has the most extensive coverage by direct output measures. These relate in the main to individual services, but, in the case of the Police, experimental work has been undertaken on a collective service. The Eurostat guidelines indicate that all functions should be covered by direct measures if an A grading is to be achieved, but input measures satisfying certain criteria are accepted as B methods for collective services.
- 5.18** Given the Eurostat guidelines, ONS should ultimately seek to extend direct measures throughout the range of government activity. However, a number of those we have consulted have urged caution. Not only are some other EU member states being slow to implement the guidelines, but also in some cases the delay is due to their belief that national accounts should have remained with an (output = input) convention.

- 5.19** In our view, **Principle C**, formal criteria should be set in place for the further extension of direct output measures. 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 allowance for quality change, (iii) the effects of its introduction have been tested service by service, and (iv) the context in which they will be published has been fully assessed, in particular the implied productivity estimate.

#### Coverage geographically

- 5.20** As noted above, the present measures differ in their coverage across the United Kingdom. In the light of devolution, and the increasing emergence of differences in services and methods of service delivery, it is important there be full use of measures from Scotland, Wales and Northern Ireland. **Principle D:** 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.

#### Inputs and Productivity

- 5.21** Government output, like private sector output, is produced using factors of production: capital, labour, energy, purchased materials and services. Collapsing the last three into ‘goods and services’, we have a three-factor production function. As noted in discussion of the OECD Productivity Manual, there may be interest in single factor measures of productivity, such as output per worker, but our primary interest is in total factor productivity.

Earlier, we have stressed the need to consider not only the output side of the equation but also the input side.

- 5.22** **Interim Conclusion (2) Regarding Inputs:** the consideration of the measurement of inputs, and of the price deflators applied, has turned out to be an important element of the review; and will form a significant part of our work in the second stage.

- 5.23** We therefore append two principles for the measurement of inputs. The first concerns coverage. Given that our interest is in total factor productivity, the inputs taken into account should be as extensive as possible in their coverage. The inputs currently considered by ONS include all three of those listed above: capital, labour, and goods and services. However, the capital element is limited to capital consumption. Capital consumption refers to the depreciation of the fixed capital, not to the opportunity cost of the capital being employed. For any given type of asset, there is a flow of productive services from the cumulative stock of past investments. This flow of productive services is called the *capital services* of an asset type and is the appropriate measure of capital input for production and productivity analysis. Conceptually, capital services reflect a quantity, or physical concept, not to be confused with the value, or price concept of capital. To illustrate, take the example of an office building. Service flows of an office building are the protection against rain, the comfort and storage services that the building provides to personnel during a given period. Because flows of the quantity of capital services are not usually directly observable, they have to be approximated by assuming that the service flows are in proportion to the stock of assets.

- 5.24** We therefore put forward **Principle E**: the measurement of inputs should be as comprehensive as possible and, in particular, should include capital services; consideration should be given to the split between current and capital spending (see paragraphs 5.26-5.28).
- 5.25** Capital services can be included in a variety of different ways, and the implementation of this recommendation will involve agreement on an appropriate convention. In considering such a convention, it will be useful to draw on the relation with the determination of the social discount rate to be applied in the ex ante appraisal of public projects (see, for example, Drèze and Stern, 1987, paragraph 3.5.2). While it may be tempting to attribute a cost of capital related to the cost of government borrowing, it is not clear that this would be correct on either theoretical or practical grounds. The valuation of capital services should certainly not be affected by short-term changes in the sources of government financing.

#### Current and capital

- 5.26** In considering capital services, we have to consider the definition of capital formation. Under this heading there are two distinct issues. The first is the line drawn between consumption and gross capital formation as a matter of accounting practice. It has been suggested to us that the present practice tends to include under 'current' items that in other countries are partly or wholly capitalised. Software is one example. A study by the OECD (Ahmad, 2003) points this up clearly. The UK practice evidently results in only around one seventh as much of total software expenditures being capitalised as in Denmark, and similarly lower proportions as compared with other developed countries, a variation not easily explained by differences between these economies. Other borderline cases, such as practice in allocating between maintenance expenditures (classified as current spending) and refurbishment (capital spending) could also stand further scrutiny. To the extent that this tends to overstate current inputs, a period of accelerated spending on these items would show up as a reduction in productivity.
- 5.27** The second issue concerns the timing of inputs and outputs. There are a number of items which may be correctly treated as final consumption but which contribute to activities at a later date. There may be lags between inputs and outputs. This applies particularly to government output such as in Education and Health. The acquisition of knowledge, skills and qualifications increases the productive potential of the individuals concerned and is a source of future economic benefit to them. (SNA 1993, paragraph 1.52). Equally, the consumption of health services brings long-term benefits.

**5.28** One simple explanation of the pattern observed in the United Kingdom is that there is a delay before the increased inputs feed through into increased outputs, or that there are once-for-all costs of expansion that will not recur. This suggests that our interpretation of the productivity record would be aided by a disaggregation of inputs. It would for example be useful to separate out training expenditure. More generally, we may wish to consider the introduction of measures of human capital formation. The SNA argues that ‘while knowledge, skills and qualifications are clearly assets in a broad sense of the term, they cannot be equated with fixed assets as understood in the System.’ (SNA 1993, paragraph 1.52). It would, however, be possible and useful to create satellite accounts for human capital formation (see paragraphs 6.21-6.25).

### Price deflators

**5.29** An essential feature of the productivity calculation is that it depends on the volume of inputs, not their value. It is here that the price indices for inputs enter the picture. For each of the three factors (labour, capital and goods and services), we need price indices appropriate to each of the spending functions. Each of these poses problems.

**5.30** In the case of goods and services, the requirements may be taken over from the *Eurostat Handbook on Price and Volume Measures in National Accounts*. Specifically, to constitute an A method, it is required that (i) there be complete coverage, (ii) the prices should be purchasers’ prices, (iii) quality change be taken into account, and (iv) there be consistency between the indicator and national accounts concepts. To these criteria, we would add the requirement that they should be sufficiently disaggregated to allow for changes in the input mix. Policy changes, such as the introduction of classroom assistants, or of community support officers, should be accommodated by the price deflators.

**5.31** In the case of labour, the problems of indices for earnings are well known. Arguing by analogy with the Eurostat Handbook requirements for the prices of goods and services, but leaving out quality change, we would expect that (i) there be complete coverage, (ii) the index should relate to actual earnings and not just wage settlements, (iii) the wage indicator should reflect total employment cost, including National Insurance and pension contributions, and (iv) there be consistency between the indicator and national accounts concepts.

**5.32** To summarise on price deflators, **Principle F** is that criteria should be established for the quality of 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.

## Triangulation

**5.33** As noted in Chapter 2, the direct approach to the measurement of government output yields an *implied* measure of ‘government productivity’. It is a residual. However, there is clearly a risk that the residual will behave in unexpected ways and that it will be dominated by the vagaries of the two measured variables. This implied measure may or may not be consistent with independent evidence on the productivity performance of the public sector. In this section, we point to the need to obtain independent evidence on productivity, as part of a process of ‘triangulation’. The independent evidence may be partial in its coverage or based on small-scale surveys, and would be weighted accordingly, but it should be part of the picture.

$$\text{Productivity} = \text{Output/Input} \quad (5.1)$$

**5.34** In terms of equation 5.1 we would be seeking to combine evidence on all three elements, rather than assuming that productivity is simply the solution of this equation. This approach would be in the tradition of National Accounts, where information from different sources has to be reconciled. This takes us back to the relation with microeconomic measures of public sector performance. As argued in the previous chapter, the National Accounts measures are not playing the same function, but they need to be coherent with the evidence from performance studies. Moreover, we need to take account of the fact, discussed in the previous section, that current spending may represent in part an investment for the future.

**5.35** This leads us to propose **Principle G**: independent corroborative evidence should be sought on government productivity, as part of a process of ‘triangulation’, taking account of the timing of inputs and outputs.

**5.36** This process of triangulation can be conducted at different levels, involving different levels of statistical resources. How far it should be taken is therefore a decision that depends on the availability of resources. Here we distinguish three levels at which the issue could be approached. In view of the importance of the issue, our preliminary position favours the most extensive programme, but we welcome comment on this position, particularly regarding the third and most ambitious level.

**5.37** As a *minimum*, the production process for the statistics of government output should involve a stage at which they are examined for coherence with other evidence. ‘Looking at the data’, along the lines of Chapter 2 above, clearly forms an essential part of the existing data production process, but this may need to be made more systematic. To draw an analogy with existing practice, if ONS finds that its statistics imply a sizeable reduction in the productivity of the mining industry, then this is presumably discussed with industry representatives to see if the finding is consistent with industry perceptions. An explanation should be provided for any divergence.

**5.38** In the present context, such cross-checking would involve consideration of the relation of the implied productivity estimates to departmental indicators of performance (including the Public Service Agreement (PSA) targets). As stressed above, these are not measuring the same thing, and there is no necessary reason why there should be agreement. It is, however, necessary that the reasons for any difference should be

understood, particularly when the direction of change is different. It is possible, for example, that departmental performance indicators cover only part of the activities, and that resources have been transferred to ensure their fulfilment. The activities not recorded in the performance indicators may have seen a lower growth, causing an overall measure of output to show a lower rate of increase than inputs. Or it may be that the PSA targets are directed at quality improvements, and that these are not adequately captured in the output measure – see the next section. The cross-checking may also reveal areas of the process that need further checking and refinement. The end product of such cross-checking would be, we envisage, a number of sentences of explanation accompanying the publication of the productivity analysis. Equally, reference has been made to the relation between the indices of labour input and the employment figures obtained from the Labour Force Survey or other sources. Such comparisons are undoubtedly made within ONS, but it would be helpful to the user if this kind of coherence check could be made more overt.

- 5.39** A *second level* would be an explicit attempt to relate the output and input indicators to departmental performance measures. From the side of ONS, this would mean a systematic examination of the relation between the direct output indicators and the PSA targets. In some cases it will be clear that the PSA targets are concerned with total outcomes (such as reducing mortality rates from the major killer diseases), whereas the national income measure is concerned with the incremental contribution of government activities. These may well be moving in opposite directions for understandable reasons. From the side of departments, evidence should be assembled regarding productivity increase. Together, these actions should increase the degree of understanding of the relation between these two different sets of indicators.
- 5.40** The first and second levels could be encompassed within the regular ‘productivity articles’ that ONS has, we understand, already decided to introduce. These articles could discuss measures of productivity in a particular sector of government output and could publish productivity estimates derived from the National Accounts with interpretation and commentary on the underlying data and issues. This commentary could have regard to the various relevant considerations discussed above. The articles could also bring together ‘triangulation’ evidence on productivity in the sector from different sources which may not be appropriate for the National Accounts but help clarify the efficiency with which inputs are being used, in the context of the main policy and delivery priorities.
- 5.41** The *third, and most ambitious*, level, would be to initiate a government productivity measurement programme. This could draw on the US experience with the Federal Productivity Measurement Program conducted by the Bureau of Labor Statistics (BLS). This was brought to a close in 1994, but statistics produced covered a period as long as 27 years. The reasons cited for its closure were ‘budgetary constraints’ (Fisk and Forte, 1997, pp 19). This serves as a warning as to the resource costs, but does not indicate that the approach as such was a failure. As described in Chapter 3, the programme used a large number of output indicators (more than 2,500 in 1994) to generate an output measure, disaggregated by functions, such as social services and benefits, or legal and judicial activities. But it would not be necessary to collect

anything like this number of indicators to achieve a major improvement in the information available regarding public sector performance.

- 5.42** On the input side, the programme collected indicators of the number of full-time equivalent employees (not adjusted for skill or experience) and of total employee compensation. From this, the BLS estimated indices for output per employee year (a 34.3 per cent increase between 1967 and 1994) and for unit labour cost. Two major limitations from our standpoint are that labour was the only input taken into consideration and that no allowance was made for changes in the skill composition of the labour force. It would, however, be possible, with sufficient resources, to overcome these shortcomings in a new government productivity measurement programme.

### Quality Change and Complementarity between Public and Private Output

- 5.43** As was observed in *Economic Trends*, July 2003, pp 27, ‘the output measures may have failed to reflect all the quality improvements made in the outputs as a result of rising consumer expectations and the more demanding standards set for service delivery’. We have seen that addressing the issue of quality change is a major concern to statistical offices in other countries.
- 5.44** We have found it helpful to make use of the framework set out by Dawson et al (2004) of the Centre for Health Economics / NIESR / NPCRDC project on measuring productivity in the UK health care sector. They consider separately the outcome per unit of activity and the value of a unit of outcome. The total effect is the product of the two elements.

#### Outcome per unit of activity

- 5.45** In the case of Health, which we take here for illustrative purposes, quality improvements may take the form of shorter waiting times, better patient experience, improved survival rates, improved quality of life, reduced re-admission rates. Some, but, not all of these quality improvements may be captured by a cost-weighted index of activities. Suppose that there are two treatments. One is of a higher quality and is more expensive. If the output indicator combines the two volume measures with weights according to their cost, then a shift towards increased use of the higher quality treatment will be properly recorded. There will be an increase in expenditure, and a corresponding increase in output. Thus, the move to a more detailed treatment classification (see Chapter 7) is a step towards taking account of this kind of quality change.
- 5.46** But, a cost-weighted index cannot capture all quality improvements, as may be seen from the example of an improvement in technique such that a lower cost treatment (eg day surgery) can assure the same (or better) outcome as a more expensive treatment. This is recorded as a reduction in output, when no change has occurred (and there has been a saving in inputs). One approach to this problem is to weight activities not by cost but by an indicator of the increment in quality of outcome, such as Quality Adjusted Life Years. In Education, the same role may be played by performance on tests. The exploration of these methods, and a testing of their validity, forms part of the ongoing research programme.

**5.47** Evidence about outcome per unit of activity can be used in two different ways. The first is to choose a particular indicator, or set of indicators, measured regularly, and to apply these on a formulaic basis in making quality adjustments. As discussed in Chapter 8, the quality of education could be made a function of the annual results in pupil attainment. An improvement in key stage results, relative to the expected transitions, would be translated into an increase in productivity. The second approach is to employ a range of evidence to make a judgement about the rate of productivity increase. This approach has the disadvantage that it involves judgements being made *ex post* (rather than in the *ex ante* selection of indicators), but the advantage that it is less at risk for example, in this specific case, from changes in examination structures or school inspection procedures.

### Valuation

**5.48** The valuation of outcome changes is an issue that arises with the quality adjustments just considered. The literature on Quality Adjusted Life Years has considered the financial value to be attached, and how it should be adjusted over time (the answer given by Gravelle and Smith (2001) is that it should grow at approximately 1.5 per cent per year in real terms).

**5.49** In considering this issue, it is helpful to distinguish between the *ex ante* and *ex post* valuation of activities. This distinction has already been noted in the previous chapter, when discussing the link with studies of the incidence of benefits on individual households. In a market context, many items take the form of insurance purchased *ex ante*, and the consumer valuation is on that basis.

**5.50** By taking an *ex ante* approach, one is in effect attaching an ‘option value’ to the existence of a public service. There is a value to a person even if he or she does not actually make use of the service. The notion is most commonly applied to natural resources. But it is the same as in market-based insurance transactions. A person taking out a gas repair insurance is securing an option on the services of a plumber. A second example is home insurance, for which people are willing to pay *ex ante*, even though most hope not to receive any payment *ex post*. The premiums paid will reflect the probability of, say, fire damage. The same may be applied to a range of public sector outputs. The most obvious example is indeed the fire service, not least for historical reasons (as revealed in the still existing fire marks). The present output measure is an *ex post* calculation of the actual number of primary and secondary fires attended (plus an allowance for fire prevention and special services). This could, however, be replaced by an expected number of fires.

**5.51** Would this make any difference? In the home insurance case, leaving aside the administration costs and the profits of insurance companies, the sum of *ex ante* premia should equal the *ex post* outlays. In expectation, the total should be unchanged. However, the year-to-year pattern will be different. In the case of the fire service, government output as at present measured varies with the number of fires. A dry summer, or a spate of arson, will raise recorded output.

- 5.52** Moreover, the premium paid for home insurance reflects both the probability of loss and the magnitude of the loss. The direct output measure in effect takes account of the former but not the latter. This neglect of the extent of property protected does not seem reasonable. (We have in mind here the replacement cost of the property, not the total cost that includes the value of the site.) The ‘benefit’ of saving a house from a fire must have gone up with rising housing standards and increased household contents. The same applies to industrial and commercial property. The prison service provides a second example. The value to society may be seen in terms of the *ex ante* capacity to incarcerate people if the need arises. The value depends in part on the scale of property to be protected and of the economic activity that would be disrupted by criminal activity. The advent of Information Technology, for example, has both increased real GDP and increased the need for protection.
- 5.53** It should be stressed that we are talking here about the real value, not the increase due to inflation. Moreover, we are talking about the marginal valuation, not about consumer surplus. We are not seeking to attribute to the fire service the full benefit from its existence. We are not measuring the area under the demand curve. Rather, we are adding the marginal willingness to pay of different citizens. Two centuries ago, when the typical home contained little of value, the value of the fire service or of police protection or the courts in terms of domestic property was quite limited. Today, with homes resembling offices/factories, and with offices and factories full of expensive equipment, the value is many times higher. To this it may be responded that the typical person two centuries ago may have owned very little but he or she could much less afford the loss. This is true, but the fact that £1 meant so much more applies also to private output, to which we are adding public output to form GDP. The person could equally not afford to lose a week’s wages. Finally, we are not suggesting that the value of public sector output necessarily increases proportionately with the value of private sector output. If people chose to spend all of their increased real income on foreign travel, then there would be no increase in the value of home contents (although there would probably be more airports).
- 5.54** The point made here is a quite general one: to a significant degree, the output of government services rises with the real value of private assets and incomes. We have just seen the force of this argument in the case of the protective services and it can be extended to health care and to education. People with today’s higher real incomes are willing to pay more to have the option of receiving medical treatment without delays. In terms of days of work lost through sickness, the valuation almost self-evidently rises with the real wage rate, so that again the valuation of the output (rather than the activity) should reflect rising real output per head in the private sector. If we see the output of education in terms of the acquisition of skills and qualifications, then their value increases with rising real earnings. If a university degree adds, say, 20 per cent to earnings, then today’s degree adds 20 per cent of a larger number (even adjusted for inflation) than the degrees of a generation ago. It should be stressed that these considerations apply to *changes* in the valuation over time. It is the value of output relative to a generation ago that is affected. Moreover, if the increase in the number of people with degrees leads to a fall in the earnings premium for graduates, then this fall has to be taken into account.

- 5.55** Unless account is taken of this complementarity, there is a serious risk that the output of the public sector will be understated. The existing direct output measures may capture increased activities but not the increased real value of those activities.
- 5.56** These considerations are summarised in **Principle H**: value should be seen as adjusted for quality; for each service, explicit consideration should be given to the incorporation of quality change as an element of value added; for each spending function, consideration should be given to the extent to which quality change is captured by the changing activity mix, and to the way in which output measures for government should be adjusted for increased real value in an economy with rising real GDP.

### Summary of Principles

- 5.57** The principles are:

**Principle A:** the objective should be to measure output from government spending in terms of its incremental contribution to individual or collective welfare, in the same way as market output.

**Principle B:** the procedure of defining direct output indicators within a government function should start by seeking to identify the services provided by government to households and firms, and attempts made to find data to reflect these services as comprehensively as possible (rather than working back from available indicators). Where, initially, it is necessary to apply an indicator from another service, this should be explicit. The coverage of indicators within a function should be reassessed on a regular basis.

**Principle C:** formal criteria should be set in place for the further extension of direct output measures. 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 allowance for quality change, (iii) the effects of its introduction has been tested service by service, and (iv) the context in which they will be published has been fully assessed, in particular the implied productivity estimate.

**Principle D:** measures should cover the whole of the United Kingdom; where systems for delivering public services 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 E:** the measurement of inputs should be as comprehensive as possible, and in particular, should include capital services; consideration should be given to the split between current and capital spending.

**Principle F:** criteria should be established for the quality of 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 G:** independent corroborative evidence should be sought on government productivity, as part of a process of ‘triangulation’, taking account of the timing of inputs and outputs.

**Principle H:** value should be seen as adjusted for quality; for each service, explicit consideration should be given to the incorporation of quality change as an element of value added; for each spending function, consideration should be given to the extent to which quality change is captured by the changing activity mix, and to the way in which output measures for government should be adjusted for increased real value in an economy with rising real GDP.

# 6 Implementation

---

**6.1** In this chapter, we consider the process of implementing the principles described above with new or different measures of government output. It deals with implementation in general, illustrating the issues by reference to specific spending areas. Individual spending areas are covered in depth in the four chapters that follow. Its role is in part to provide a bridge to those chapters. Whereas in the previous chapter, we considered the issues theoretically, starting from the desired objective and seeking to derive principles to be applied in the construction of output indicators, the present chapter works in the opposite direction. It starts from existing ONS practice and examines how we can move towards implementing these principles more fully. It does so on the basis that the move to direct output measures is justified both by its intrinsic merits and by the obligations placed on the United Kingdom by the Eurostat procedures.

## Output Measures

**6.2** Implementation of the principles set out in Chapter 5 means that we have to consider separately:

- what changes, if any, should be made in those spending areas where direct measures have already been introduced;
- whether direct output measurement should be extended to new spending areas; and
- if an input basis is retained, how the estimates could and should be improved.

**6.3** The direct output measures introduced by ONS over the period 1998-2001 represented a pioneering ‘first generation’ set of methods. As improvements become possible, we would expect ONS to replace them by more refined ‘second generation’ methods. As in other fields, the development of national accounting techniques is a dynamic process. Current direct measures of output can be improved by:

- a) widening the coverage of output volume indicators for each function;
- b) increasing the level of detail at which output indicators are measured (and the corresponding weights);
- c) replacing activity indicators with output measures that reflect changes in quality or outcome attributable to a unit of output;
- d) improving in-year indicators; and
- e) improving UK coverage by making full use of measures from Scotland, Wales and Northern Ireland.

- 6.4** In order to introduce a replacement output measure, the following conditions should be satisfied:
- a) There should be evidence of improvement in one or more of the five directions listed above, giving particular emphasis to completeness of coverage and to measures that reflect quality change.
  - b) An analysis should have been carried out of the relevant output, input and price deflator data from past years, with sensitivity testing for possible future changes, to assess the direct output measure and the implied productivity measure.
  - c) The validity of the proposed measure should be tested by those with expert knowledge of the relevant function.
  - d) There is assurance of the likely continuation of the key data sources which are ideally drawn from data used for management and operational systems.
- 6.5** The implications for different spending functions are discussed further in Chapters 7 to 10. It is clear that the approach to be taken in different areas depends on whether the function as a whole provides collective or individual services. Guidance on which services should be regarded as falling in each category is limited and open to debate. We need to give more consideration to the classification of different functions overall, and parts of services. For example, Health is generally an individual service but includes public health functions which are collective services, whose output can only be measured as such. Most of the Public Order and Safety function can be regarded as a collective service, but it includes functions like victim support which are individual services whose outputs should be measured appropriately.
- 6.6** The second issue concerns the extension of direct output measures to functions not already covered. We recognise the intrinsic merits of a direct measurement approach, but believe that ONS should proceed circumspectly. A number of those we have consulted have urged caution and we believe this counsel to be wise. Caution does not, however, equate with neglecting this agenda. Prudence in decision-making is fully consistent with the continuing vigorous work programme we believe to be necessary.
- 6.7** A distinction should be made between individual and collective services. In the former case, the European Commission Decision means that ONS is required to extend direct measures throughout the range of government activity by 2006. Given that the direct measures already cover two-thirds of total government output, the United Kingdom appears well placed. We recommend, nevertheless, that ONS determine, in consultation with Eurostat, which of the remaining functions should be classified, either in part or whole, as individual services.
- 6.8** In the case of collective services, it will remain possible to continue to use an (output = input) approach (as a B measure). But it is not clear that the existing measures will be regarded as satisfactory. There are two reasons for this: because of the quality of the input data (considered in the next section) and the Eurostat requirement that appropriate account be taken of quality changes. Further work will be required in this area.

**6.9 Interim Conclusion (3) Regarding the Extension of Direct Output Measures:** The overall thrust of the Commission Decision is welcome and we believe that ONS, like other statistical offices, should work vigorously on this agenda. Implementation decisions, however, should be taken with due caution against a clear knowledge of the implications and against clear criteria.

### Input Measures and Deflators

**6.10** Our examination of the provision of government data to ONS has identified a number of serious issues. As noted in Chapter 2, the process is a complex one. The data required to compile the estimates has at least the following seven sources:

- central government departments via the Treasury;
- English local authorities via ODPM;
- Scottish local authorities via the Scottish Executive;
- Welsh local authorities via the Welsh Assembly ;
- Northern Irish local authorities via Northern Ireland Statistics and Research Agency data via the Treasury;
- Customs and Excise; and
- Inland Revenue.

Although it was not their primary focus, the report by McLean et al (2003) on regional public spending flows identified very clearly the complexity of the data flows.

**6.11** The lack of clarity about the overall process, as identified in Chapter 2, contributes to a number of problems including:

- a poor match of data breakdown to National Accounts needs, requiring statisticians to make a number of assumptions to enable the compilation of components;
- an apparent inconsistency of the capital consumption data with data provided by the Treasury and ODPM;
- reliance on a relatively small number of pay and price deflators of questionable quality; and
- the relationship between the labour input data and the public sector employment data needs to be reviewed.

**6.12** These issues raise problems both for the implicit measure of productivity and for the continued use of an (output = input) approach for collective services. They need to be tackled.

**6.13** The issues identified above can be traced in part to the absence of service level agreements. In our view, there is insufficient recognition across government of ONS as an important data user and of the importance of accurate National Accounts. In turn, ONS needs to develop stronger linkages with its data suppliers.

### New data sources

- 6.14** New sources of audited financial year data are gradually becoming available through the ‘dry run’ processes leading to the publication of accounts based on Generally Accepted Accounting Practice – the ‘Whole of Government Accounts’ – covering first central government and subsequently the whole of government. These audited outturn financial year data will also be integrated with the quarterly central government data and other financial year data produced by the Treasury through its Single Data System (SDS) project, which is due to be implemented during 2005/06.

### Timing and timeliness of data

- 6.15** From the analysis of the data flows under the arrangements in place when this review began, it appears that the following problems occur:

- Much of the data relates to financial years, and has to be converted to a calendar-year basis for National Accounts purposes.
- Quarterly data are necessary to make this adjustment, and are necessary in their own right to produce quarterly GDP estimates. Quarterly GDP estimates are of lower quality than annual figures and make use of interpolation. (See *Economic Trends* November 2001).
- There are considerable, in some cases, astonishing, delays in the supply of financial year data; although the moves towards faster closure of accounts in both central and local government should help reduce these over the next few years.

- 6.16** As a result, the most recent estimates are subject to significant revision, and the quarterly estimates are of lower quality than the annual data. It seems increasingly likely that in-year financial management data are used by local authorities. It would be worrying, from the point of view of the stewardship of public funds, if they were not, and further investigation is being carried out in this area to determine the scope for making use of this data.

- 6.17** The devolved administrations supply data on their local authorities. Data supplied are only annual, on a financial year basis. There are no formal agreements for data supply between the devolved administrations and ONS. Improving the quality of data from devolved administrations would make a major contribution to increasing UK coverage.

- 6.18** In order to make progress with these issues, we recommend **Interim Conclusion (4) Regarding Input Data:**

- a) ONS, the review team and the Treasury should work together to review the data supply and the processing chain for central government estimates, to feed into the Treasury review already underway.
- b) ONS, the review team and ODPM should agree an action plan to address data timeliness, breakdown, and periodicity of local authority data.

- c) ONS and the review team, with the assistance of the Treasury and ODPM, should carry out a study to establish whether ONS should replace the capital consumption estimates modelled by ONS with the capital consumption estimates supplied by the Treasury and ODPM.
- d) ONS and the review team, together with relevant departments, should carry out a review of pay and price deflators.

### Further Consideration

- 6.19** In Chapters 1-5 a number of ideas were introduced that either go beyond existing National Accounts practice or are speculative in nature. They are put forward to provoke further thought and examination, rather than for immediate implementation.

#### Capital services

- 6.20** The first can be treated briefly here, as it was extensively discussed in Chapter 5. This is the introduction of capital services. We are not proposing that this be introduced immediately (although the neglect of capital services may be part of the explanation for the behaviour of the implied productivity series). We would strongly encourage its introduction in the medium-term, and urge that consultations be undertaken with the SNA Revision Advisory Expert Group and with Eurostat.

#### Satellite accounts

- 6.21** In the context of human capital formation, it was suggested there may be a role for satellite accounts. The ESA notes that ‘for some specific data needs the best solution is to draw up separate satellite accounts ... Satellite accounts can serve such data needs by:

- showing more detail where necessary and leaving out superfluous detail;
- enlarging the scope of the accounting framework by adding non-monetary information, eg on pollution and environmental assets; and
- changing some basic concepts, eg by enlarging the concept of capital formation by amount of the expenditure on research & development or the expenditure on education.’ (ESA, 1995, paragraph 1.18 and 1.19).

- 6.22** It is important that satellite accounts be consistent with national accounts principles. The ESA states that, ‘an important feature of satellite accounts is that in principle all basic concepts and classifications of the standard framework are retained. Only when the specific purpose of a satellite account definitely requires a modification, are changes in concepts introduced. In such instances, the satellite account should also contain a table showing the link between the major aggregates in the satellite account and those in the standard framework. In this way, the standard framework retains its role as a framework of reference and at the same time justice is done to more specific needs.’ (ESA, 1995, paragraph 1.20).

- 6.23** A number of countries have developed satellite accounts. INSEE, the French national statistical institute, for example, has developed satellite accounts for Education, the Environment, Tourism and Health. It has also been at the forefront of the development of satellite analyses, which are similar in nature but are less closely intertwined with the main National Accounts, allowing for quicker delivery and greater flexibility.
- 6.24** The United Kingdom currently produces satellite accounts for three areas (Environment, Households, and Health) and is developing prototype satellite accounts for Tourism. It is aware of interest in the potential for development of satellite accounts for Transport and for Culture.
- 6.25** The relationship between Satellite Accounts and National Accounts may also be useful in providing a model for the relationship between estimates of government output within the National Accounts framework and productivity estimates produced outside of the National Accounts framework. We have also suggested in paragraph 5.28 that it would be useful to have a satellite account for human capital formation.

#### Complementarity between public and private output

- 6.26** In Chapter 5, it was argued that the value of the public sector output may rise with the real growth of the private economy where the public sector activity complements that of the private sector. Although the idea is present at certain points in the professional economics literature, it has yet to be widely accepted. It is put forward here to generate further discussion.

#### Process of Implementation

- 6.27** There are several important elements in the process of implementation including:
- agreement with National Accounts Group, and other units in ONS;
  - action plans agreed with departments to provide co-operation and ownership of work done during the lifetime of the review and beyond, and to clarify flows of data from departments to ONS; and
  - consultation with the Statistics Commission and other external bodies.
- 6.28** ONS naturally already has means of quality assuring changes to methods in the National Accounts. The current process is:
- a) a peer group appraisal for the conceptual basis for the changes, to which external reviewers are invited;
  - b) discussion of the outcome of peer group appraisal by the National Accounts Methods Board on whether acceptance of the conceptual framework should be recommended to the National Accounts divisional directors;
  - c) once the conceptual framework is accepted, then a second peer group appraisal to review the new estimates resulting from the changes to the conceptual framework;

- d) a discussion of the outcome of this second peer group appraisal by the National Accounts Methods Board on whether acceptance of the new method and estimates should be recommended to the National Accounts divisional directors; and
- e) inclusion of new estimates in the National Accounts framework as permitted by the revisions policy.

**6.29** Ownership of the process by all involved is essential if reliable and timely figures are to be produced. At the moment, we do not believe that the process is sufficiently transparent, nor that the suppliers of data are sufficiently aware of the importance of the figures they supply for the construction of the National Accounts. In turn, ONS needs to show greater awareness of data suppliers' circumstances. Achieving better linkages is a two-way process. The review has focused attention on this area of government statistics, and it is important that machinery be put in place to ensure a continuing process after the review has been completed.

### Timing

**6.30** The way in which revisions to existing statistics are made is of importance, as the Statistics Commission report, *Revisions to Economics Statistics* (2004), underlines. Within the constraints identified above, including the Eurostat deadline for compliance, changes to the output and input measures could be introduced, a) progressively, as work on each service or function is completed; b) in several blocks, grouping changes; or c) in one single change. The last of these could mean considerable delay, and that we would be continuing to use measures that are less than fully satisfactory for a significant period after superior measures became available. The first would lead to a series of revisions which could prove confusing for users. We therefore favour the second approach: a balanced compromise between delaying improvements in estimate quality and repeatedly revising estimates.

### Future review process

**6.31** Consideration needs to be given to the future review process as well. In our view, there is a good case for a regular review of the statistics of government output and productivity. The shape and policies of the public sector change with changing political priorities. Changes in the organisation of public services were one reason why the present review has been necessary. There is every reason to expect that organisational and technological change will be just as rapid in the future as over the past decade.

### Publication and Presentation of Estimates

**6.32** The public attention recently focused on the National Accounts measures of government output and productivity has lead us to consider their public presentation.

### Publication

**6.33** ONS is to be congratulated on the extent to which it has published the new estimates and the underlying methodology for government output. The series of *Economic Trends* articles has performed a valuable service, and it seems to us important that

ONS should continue to provide estimates in the detail exemplified by the July 2003 *Economic Trends* article. This includes the series for the implied index of productivity, described as ‘experimental’. Progress towards better measures of output and inputs will be aided if the estimates are exposed to public scrutiny.

- 6.34** At the same time, there is a risk that the National Accounts measures of government output and productivity will be misunderstood. Two ways in which this risk can be reduced are by (a) reiterating clearly the difference between National Accounts aggregates and micro measures of performance and (b) the proposed publication of ‘productivity articles’.

#### Margins of error

- 6.35** Consideration of the estimates for government output, inputs and productivity raises a more general question about the margins of error that surround these estimates. In the past, the CSO had its A, B and C classification, which corresponded to margins of error of  $\pm 3$  per cent, from  $\pm 3$  per cent to  $\pm 10$  per cent, and more than  $\pm 10$  per cent (Maurice, 1968, pp 40). These were described broadly as ‘good’, ‘fair’ and ‘poor’. These ratings were based on a subjective assessment, and have since been dropped. ONS currently uses revisions analysis as the basis for an assessment for reliability.
- 6.36** We appreciate the problems in providing an empirical basis for such ratings, but feel that the issue should be given more consideration. The lessons from the revisions analysis should be spelled out in all relevant data publications. Research in this area should be a priority.
- 6.37 Interim Conclusion (5) Regarding Presentation:** Greater consideration should be given to the presentation of National Accounts measures of government output and productivity, bringing out the purposes for which they should and should not be used. National Accounts should be coherent with, but are different from, PSA targets and micro-measures of public sector efficiency; explicit reference should be made to the margins of error surrounding estimates.

#### Internal documentation

- 6.38** The compilation of this component of GDP is extremely complex and involves the combination of inputs from a wide variety of sources, coupled with extensive use of interpolation and forecasting. It is therefore particularly important that these processes be fully documented. There is at present a serious risk that lack of clarity across important players in the data processing chain may put the reliability of the estimates at risk.
- 6.39** This applies particularly to the input side and to the price deflators. The flows of data between departments, the Treasury and ONS need to be set out clearly. The relation with accounting data used by the National Audit Office and others, and the Whole of Government Accounts, needs to be clarified in order to maximise the benefits for the National Accounts from both Whole of Government Accounts and the SDS project.

**6.40 Interim Conclusion (6) Regarding Documentation:** ONS and the review team should work closely with relevant government departments to ensure that documentation is produced that fully informs all parties, including departments and the public, about the processes involved in compiling the government expenditure component of the National Accounts.

### Resource Costs of Statistics

**6.41** In this and the previous chapter, we have sought to amplify the desiderata for effective measurement of government output and productivity. We are conscious that many of these would involve additional statistical resources and place additional demands on government departments. These demands would in part be an investment in the setting up of a new system, but there would also be substantial continuing running costs.

**6.42** Looking back, it seems clear to us that the original move to introduce direct output measures increased the demands on ONS and on departmental resources to an extent not fully appreciated at the time. It must be emphasised that additional resources will be needed to enhance the data supply process and the analysis of the data. We appreciate the current pressures on departmental staffing levels, but greater priority will have to be given to data provision activity if improvements in the measurement of government output are to be achieved. As brought out well in the *Allsopp Review of Statistics for Economic Policymaking*, good data require resources. The statistical resources allocated to measuring government output should be at least commensurate with its contribution to GDP; moreover, the areas within the scope of this review are ones where there is a particular premium on good data.

**6.43 Interim Conclusion (7) Regarding Resources:** If the government wishes to have reliable estimates of government output and productivity, then the statistical resources have to be supplied.

# 7 Health: A Progress Report

## Introduction

- 7.1** In the United Kingdom, central government provides through the National Health Service (NHS) a comprehensive health care service funded from general taxation.
- 7.2** The Department of Health (DH) in England provides funds to Primary Care Trusts (PCTs), which in turn fund NHS Trusts for providing hospital and some community health services. PCTs arrange for primary care for their local populations, and may provide some community health services directly. PCTs have an important role in planning and providing public health services including health promotion. Hospitals are operated by NHS Trusts, which were classified as public corporations until June 2004 but have now been reclassified as part of the government sector. Services are free of charge at the point of delivery except where patients are liable to pay prescription and dental charges.
- 7.3** The provision of health care services in Wales, Scotland and Northern Ireland is a devolved responsibility and there are important differences in the organisation of services in each country. These differences are not discussed further in this report, but work with the devolved administrations will be important in the next stages of our work programme.
- 7.4** Health is by far the largest of the government services as measured in terms of money turnover. In 2003, it accounted for 31 per cent of government final consumption.
- 7.5** The aim of DH, under the 2002 Public Service Agreement, is to transform the health and social care system so that it produces faster, fairer services that deliver better health, and to tackle health inequalities. The high level objectives are to improve service standards and improve health and social care outcomes for everyone. Specific targets include:
- quicker access to treatment – General Practitioner (GP) appointments, accident and emergency waiting time, outpatient appointments and inpatient admission;
  - achieving outcomes such as lower mortality in specific disease areas; and
  - improved patient experience – more choice in hospital appointments and improvements in other aspects of health care experience important to patients.
- 7.6** These are three of the domains in which ‘quality’ can be measured as part of Health output. There are others, including a reduction in morbidity through prevention of illness, or treatment and care that reduces pain, disability and dependence for those who are ill, including people with chronic diseases. The need to measure some or all of these aspects of quality as well as the volume of health care activity is discussed further in paragraphs 7.32 to 7.38.

- 7.7** DH also has targets to reduce health inequality. Distributional issues are not addressed in National Accounts measures, but they are important in a wider assessment of NHS productivity performance.

### How UK Health Outputs are Currently Measured

- 7.8** Following international guidance, the UK Health output measure tries to reflect the volume of goods and services produced as health care outputs. Health care outputs are those which benefit or increase the welfare of recipients. Most of these are individual recipients receiving benefits such as hospital treatment, a GP prescription or a visit from a district nurse.
- 7.9** The UK Health output measure used until June 2004 reflected movements in 16 different activity series measuring health care in England. These series ranged from some that accounted for significant amounts of expenditure – such as a count of inpatient and day cases – to health visitor and district nurse episodes. All the series used in the calculation were derived from NHS operational data systems, except the measure of general practitioner consultations (including home visits, telephone consultations and practice nurse consultations), which was taken from the General Household Survey (GHS).
- 7.10** An aggregate index was formed from these series by weighting them together to reflect the amount spent on each. This index was therefore heavily influenced by movements in the count of inpatients and day cases, which together accounted for over half the spending on the activities included in the index. Only annual figures were available, after a substantial delay: *Blue Book 2003* used figures from 2001/02. Consistent data series on health care in Wales, Scotland and Northern Ireland, (18 per cent of government health care spending in the United Kingdom), are limited so their contribution to the total includes a significant amount of estimation.
- 7.11** This approach to measuring Health output would seem to rank as a C method in Eurostat's scheme as described earlier and hence would not meet the EU's standard for compiling National Accounts from 2006 onwards. This is because there is too little scope for changes in the mix of outputs – for example, a quality improvement resulting from a shift to more expensive types of treatment – to be reflected in the aggregate index. In the method used to June 2004, carrying out a heart operation adds the same amount to the index as does a much simpler and cheaper treatment.
- 7.12** As a result of work carried out by the review and ONS National Accountants, in close co-operation with Department of Health officials, an improved version of this methodology was introduced into the National Accounts calculations published on 30 June 2004. It uses Health Resource Groups (similar to Diagnosis Related Groups used internationally) and the national schedule of reference costs. The National Statistics website [www.statistics.gov.uk](http://www.statistics.gov.uk) explains the changes.

- 7.13** The improved method overcomes one main shortcoming of the previous approach: that, within each of the 16 categories, performing each type of treatment added the same amount to output. The overall output measure is now derived using a much larger number of series. In the latest year, over 1,700 different treatment types together with their unit costs and incidence are used. They range from a GP prescribed drug valued at less than £10 to a bone marrow transplant costing £99,000. As a result, changes in the mix of activities carried out will be reflected in total output proportionately to relative cost, as with most indices in the National Accounts. In principle, the more detailed method could have led to either an increase or decrease in the growth rate of measured output. The impact depends on the changes in the mix of treatments that has actually taken place over the period in question.
- 7.14** The new series also covers a wider range of NHS activities. It has added measures for the new NHS Direct and NHS walk-in centres (less than one per cent of spending), and improved the analysis of primary care prescribing by using a more detailed breakdown of prescribing activity and unit costs. Timeliness is much improved.

### International Context

- 7.15** The international context has been described in Chapter 3. Specific guidance on the measurement of health care is found in *Eurostat Handbook on Price and Volume Measures in National Accounts* (section 4.13). The Eurostat Handbook says:

‘For hospital services, output (= treatments) can be measured on the basis of so-called Diagnosis Related Groups (DRG) type classifications. DRG systems are used to classify hospital stays into groups that are medically meaningful and as homogeneous as possible with regard to resource use ... In recent years DRG systems have been introduced in many countries to assist hospital management and funding decisions. DRG systems vary across countries, but they are sufficiently similar. They are always very detailed, consisting of several hundreds of diagnosis related groups.’

- 7.16** The Handbook suggests a number of possible A methods together with their A/B/C ranking. Inputs are always considered a C method.

- For services to inpatients by general and specialised hospitals:

‘the use of fully quality-adjusted DRGs is an A method. While DRGs capture changes in the treatment mix well, changes in the quality of individual treatments are difficult to measure. They may be due to better performing equipment, better performing doctors and nurses or changes in the “hospital environment” such as the occurrence of infectious diseases in the hospital, medical errors, additional facilities for patients etc. Further research on appropriate indicators is needed.

DRGs that cover only changes in the treatment mix will fulfil the requirements for a B method ...

Use of crude output indicators like the simple number of discharges is classified as a C method.’

- For medical practice services (general practitioners)

‘the services of GPs are such that each visit can be considered as constituting one treatment. Consequently, the recommended A method is the number of consultations by type of treatment, adjusted for changes in quality. It might be difficult, however, to obtain the corresponding cost weights. In the case of proxy weights or only partial quality adjustment the number of consultations by treatment is a B method. The simple number of consultations can also be accepted as a B method if the different types of treatment are sufficiently homogeneous with regard to the resource requirements (similar cost weights).’

- 7.17** An acceptable measure of Health output, according to the most detailed national accounting guidelines, is one that measures the volume of a large number of treatments and adds each of them together, using cost weights to provide an overall total. This does not capture changes in the quality of each treatment directly. But since it allows for a higher priced treatment to add more to the overall volume than a lower priced one, it is deemed as capturing some of the increase in the quality received by patients.
- 7.18** The Netherlands, Italy and Finland are in an advanced state of preparation for the Eurostat requirements. These countries typically monitor the volume of nearly 1,000 different treatments and diagnostic activities. In the case of the Netherlands, account is taken not only of the type of treatment but also of the age group of the patients. This allows the different cost of giving a particular treatment to different age groups to be reflected in the overall total.
- 7.19** As the Eurostat guidance is the most specific international guidance, it is also seen as a useful standard by non-EU countries. Australia and New Zealand have already implemented this type of output measure, taking into account nearly 700 different treatments. Norway is preparing to use a similar approach. More detail of other countries’ experience and plans is in Appendix D paragraphs D4-D20.

### Critique of Current Health Output Methods

- 7.20** The more detailed approach introduced for the 2004 National Accounts is a significant improvement which we greatly welcome. But there is more to be done before Health output is measured accurately and completely. The 2004 method may be rated ‘B’ in Eurostat’s scheme. Weaknesses include data coverage limited to England and there are doubts about the quality of the data on general practitioner services. Moving to an ‘A’ assessment would also require use of quality measures.
- 7.21** This section examines what might be done to improve the methods further, following the framework set out in paragraph 6.3.

### Widening the coverage of output volume indicators for each function

- 7.22** The Department of Health estimates that the activity volume measures used in the new output measure cover about three-quarters of all expenditure on NHS health care activity in England measured by expenditure in 2002/03. This is based on comparing expenditure on activities which were part of the output measure in 2002/03 with NHS revenue expenditure, after excluding central budgets such as education and training expenditure which are not spent directly on health care activity. Work already underway, as a result of extension of the reference costs system, will increase coverage to almost 85 per cent in 2003/04 and even further in subsequent years. This is consistent with Principle B (see paragraph 5.16) on coverage of output measures.
- 7.23** Areas not currently covered include a wide range of activities on which expenditure is relatively small. They include continuing and intermediate care, services for people with learning and physical disabilities, national screening programmes and other public health interventions.
- 7.24** Failure to measure certain activities does not automatically lead to an understatement of the volume of output. The ONS convention is to assume the excluded activities are growing at the same rate as those which are measured. This may either understate or overstate actual change. The potential error from this assumption reduces as the proportion of all activity covered by specific measures increases. While further extension of the output measure to capture change in all distinct activities is desirable in principle, there may be diminishing returns from effort to extend coverage once it reaches about 90 per cent, and other areas of change may be more important. But sensitivity testing would help to set priorities, particularly where areas of health care activity with no direct measures are known to be priorities for expansion or reduction.
- 7.25** General medical services, the important first contact with the health service for most patients, account for 13 per cent of health care expenditure. Consultations with GPs and other members of the primary care team (eg practice nurses) are increasingly important in health care delivery. The NHS does not collect comprehensive consultation data directly from GPs. The main source available is the General Household Survey (GHS). This asks people if they have visited their GP in the last two weeks. This includes other types of contact like a telephone consultation with a GP, and checks how many occasions this occurred. There may be sampling errors although the GHS uses a 20,000 random sample of the Great Britain population.
- 7.26** Analysis of the GHS data suggests that the percentage of the population who have consulted their GP within the last two weeks has not changed significantly over the past ten years, having fluctuated in the range of 13 per cent to 16 per cent. Nor has there been change in the number of consultations per person by age band. The total number of consultations has increased as the population has increased and aged. The GHS is a useful source for data such as consultation rates for different age groups, as a snapshot for any one year. However, it does not seem to be robust as a way of measuring changes from year to year.

**7.27** There are at least two other different longitudinal sets of GP data (Royal College of General Practitioners (RCGP) spotter practices and General Practice Research Database (GPRD)), which could be alternative sources of data. Further work is needed to examine these options, or to consider other ways of collecting GP activity data directly, so as to provide good estimates of changes from year to year.

**Increasing the level of detail at which output indicators are measured (and corresponding weights)**

**7.28** The improvement in the measure introduced in June 2004 over that used previously, comes mainly from an increased level of detail. This became possible because the NHS has developed robust costs for a standard list of treatments, known as reference costs, which are being used throughout the NHS as a basis for ‘payment by results’. These allow for more detailed cost weights to match to counts of activities, which have been available for some years from the Hospital Episode Statistics (HES).

**7.29** It is unlikely that there would be any benefit in using even more detailed breakdowns of hospital activities. Ideally, better information would come from re-aggregating some measures to give a unit of output which represents the ‘care pathway’ for an illness. A patient’s health care pathway could include several linked outpatient attendances, investigations, in-patient stays where they may be transferred between consultants, and follow-up care including GP consultations and prescriptions. These will, in practice, count as several units of activity in different component parts of the Health output index.

**7.30** A change in medical practice such as extra consultations with a GP with a special interest, or undergoing investigations as an outpatient rather than receiving inpatient care – could raise or lower the total counted activities. But these changes in practice will not necessarily benefit the patient to the extent that an increase in the counted units of activity implies. If it were achievable, a unit of output based on a complete care pathway would be more meaningful.

**7.31** Such an output measure could take better account of the quality of care given. Readmission soon after treatment may be a sign that the treatment has been unsuccessful and this should ideally be flagged in the way the data are collected. (The readmission may be a ‘broken brick’, in the analogy used in paragraph 5.4, i.e. resolving problems created by clinical failure should not be counted as a new output.) NHS data systems have not in the past attempted such linkages, though more may be possible when current major IT developments are complete. We are aware that other countries are also attempting to develop systems to link relevant output episodes together. Another example of the benefits of analysing outputs differently is set out in paragraphs 7.32-7.38.

### Improving measures of output to capture changes in quality

- 7.32** The current methods capture *activities* carried out. Under Principle A (see paragraph 5.7), an ‘output’ measure should identify the incremental contribution of the activity to individual or collective welfare. This should include capturing any change in outcomes which is attributable to the use of the inputs. A basic count of activities does not measure the quality of the output, or any change in quality or clinical effectiveness.
- 7.33** The earlier discussion of the government’s stated priorities for health suggested that the main domains for understanding quality of health care are:
- a) saving lives and extending life span;
  - b) preventing illness and mitigating its impact on the quality of life;
  - c) the speed of access to treatment; and
  - d) the quality of the patient experience.
- 7.34** These are all aspects of individual welfare. Health experts might use different terms to describe these aspects of care, and we welcome views on how best to make use of expert knowledge and research in this area. We consider all these areas of quality relevant for the National Accounts and need to find ways of measuring them in conjunction with activity counts. Measuring changes over time in at least some of these aspects of quality would come closer to capturing for the National Accounts the totality of the value added by the health service.
- 7.35** There are data on changing mortality rates. These are affected by factors other than health care but research into particular conditions gives a good base for assessing the extent to which health care makes a difference. There is also a growing databank for measures of quality of life of patients after different health care interventions. Data are available on some aspects of waiting times before treatments. Patient experience is measured by individual Trusts through a national patient survey programme, using questions on five dimensions of care: access and waiting; quality and co-ordination of care; information and choice; relationships with staff; environment. Results are available at national aggregate level from the Healthcare Commission.
- 7.36** Our focus is on the year on year change in quality of health care received by individual patients, attributable to the NHS. This can sometimes be easier to measure than the absolute level of ‘quality’.
- 7.37** DH is funding the University of York Centre for Health Economics and the National Institute of Economic and Social Research to develop new approaches to measuring and understanding NHS outputs and productivity. Their work began in March 2004 and will not be complete until late 2005 but we have benefited from early discussions (see paragraph 5.44).

**7.38** As well as assessing the availability and usefulness of data in different areas, further consideration will be needed on how different aspects of quality might be combined in an aggregate measure. Health research uses Quality Adjusted Life Years (QALYs) to measure the value of a longer life in combination with different levels of pain or disability, where valuations are based on the views of those who have experienced the clinical condition. But even QALYs do not take account of other aspects of health care experience which may be important to patients, such as choice of appointment time or quality of hotel services.

#### Improving UK coverage by making full use of similar measures from Scotland, Wales and Northern Ireland

**7.39** The output estimates used in the National Accounts take data from England and gross up by expenditure weights to the United Kingdom. Based on the improved methods introduced for England in June 2004, it is now a priority to use the same or similar methods to measure activities in Wales, Scotland and Northern Ireland, based on data from their respective health care systems. The devolved administrations together account for 18 per cent of UK Health spending and could well exhibit different trends in activity and/or output.

#### Improving in-year indicators

**7.40** Estimates of GDP and its components are published by ONS on a quarterly basis. A first estimate is published about a month after the end of the quarter being measured. Two further estimates then follow at monthly intervals. These three separate estimates appear not only because there is a demand for regular updates but also because new information becomes available that users want to see included in GDP estimates as soon as possible. The introduction of the latest methodological change has advanced the date at which health care activity estimates are available. However, they are not available early enough to be used in making the first two estimates after the reporting quarter, which still have to be extrapolated from whatever information is available about health care trends, including figures on expenditure. It would be desirable to make better use of in-year information on activities or quality (which hospitals may well have for their own purposes) to improve the accuracy of quarterly updates to GDP.

### Ways Forward on Output Methods

#### Improving coverage

**7.41** We expect ONS to continue to work with DH to take advantage of the widening scope of reference costs to extend the coverage of the current activity measure. That could also include discussion on the feasibility of improving in-year estimates.

**7.42** As part of the review, we will work with DH on sensitivity testing, using any sample or local data which may be available, to consider which activities not currently measured nationally should be particular priorities for better measures.

### General practice

**7.43** We will also work with ONS experts working on the General Household Survey to consider what the error rates may be in the estimated GP consultation rates and GDP estimates based on this source. If this confirms a need for improvement, we will review with DH whether better methods can be found to measure changes in general medical services activity and/or output, taking account of the new GP contract and other structural changes. One possibility may be to use a long-term database of information about a subset of GP practices such as the RCGP or GPRD sources, though the practices involved may not be a representative sample. We are aware that the York/NIESR research study is also considering how best to measure outputs from primary care, but would like to see further work in the interim on alternatives to relying on the General Household Survey estimate as part of the National Accounts.

### Technical change, substitution, care pathways

**7.44** We do not expect to be able to make early progress in measures based on aggregating NHS episodes into ‘care pathways’. However, we would encourage further exploration of the concepts, to consider what might be done to give the clearest picture of value added by the NHS. We also consider it very important that developments in NHS information systems, while intended primarily to support patient care, should also support better future analysis of health care outputs.

**7.45** Nicola Mai, an ONS economist working on the review, has been examining a method for measuring health output using a diagnosis-based approach, which is presented in a paper published on the ONS website. The method has three key characteristics:

- the unit of measurement is the ‘whole course of treatment’ of a patient rather than activities performed;
- units of measurement are grouped by diagnosis; and
- treatments are adjusted for quality factors.

**7.46** Mai illustrates this approach with a case study on a single diagnosis – coronary heart disease. He uses a common cost weight for two substitutable treatments (coronary artery bypass graft (CABG) and percutaneous transluminal coronary angioplasty (PTCA)). The clinical pattern has been a fall in the more expensive treatment, CABG and a rise in the cheaper one, PTCA. His proposed methodology suggests higher output compared with the 2004 ONS method. If clinical opinion and research evidence is that patients have done as well or better with PTCAs as with CABG, then it is misleading to say that health output is lower simply because the cheaper treatment has been used. Where treatments are substitutes, and especially where the cheaper one gives better results for the patient, Mai’s technique has a distinct advantage. But our view is that detailed resource costs, as introduced by the ONS in 2004, are a better way of reflecting the overall pattern of change in output.

**7.47** Mai also points to further useful lines of work, particularly in combining outpatient, hospital, drug, GP consultation and GP prescribing data with hospital inpatient data to give a clearer analysis of the cost of care pathways. His methodology could be used to measure quality gain where one form of substitutable treatment has a clear advantage over another. We will consider further whether an approach which identifies benefits from technical innovation and substitution, especially if it could be based on the costs and benefits of complete courses of treatment, might be used to better understand changes in health outputs in certain areas. Until a wider range of examples are available, this could be explored as part of the triangulation approach recommended in Principle G (see paragraph 5.35).

**Measuring quality**

**7.48** We will hold further discussions with the York/NIESR research team during the remainder of the review. Although the final report in January will be published when they are only part of the way through the project, we hope to be able to say more in the final report about the most promising directions for future work from our perspective on measures used in the National Accounts.

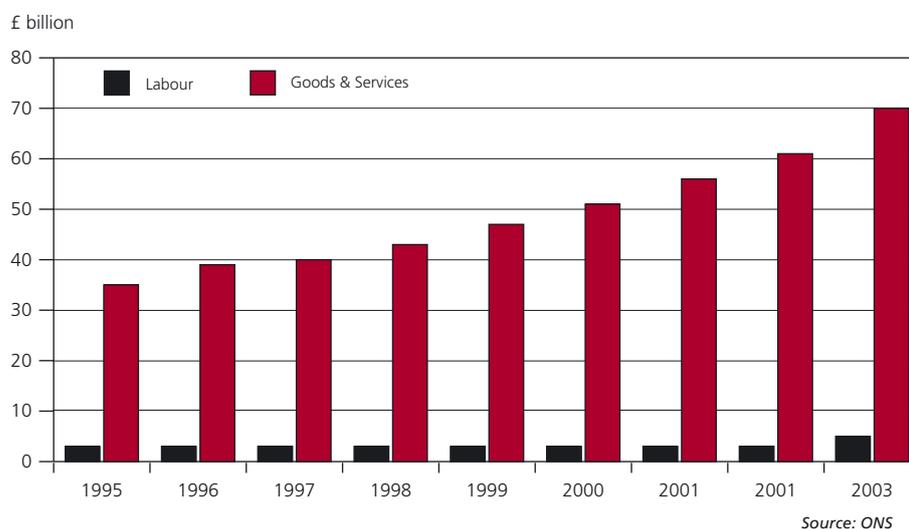
**UK coverage**

**7.49** We intend to work with ONS, DH and the health departments in Wales, Scotland and Northern Ireland to see how far relevant data from their respective health systems can be incorporated into the UK National Accounts, using the techniques adopted for England in 2004 or adapting them as appropriate, given differences in systems for delivering and recording health care, under the principles set out in this report. This would meet Principle D (see paragraph 5.20) on UK coverage. We expect to report proposals for implementation (or development work where needed) in our final report.

**Inputs and Deflators**

**7.50** The three components which make up the total expenditure on inputs in the National Accounts are labour, intermediate consumption of goods and services, and capital consumption.

Figure 7.1 **Health Care Expenditure (Current Prices)**



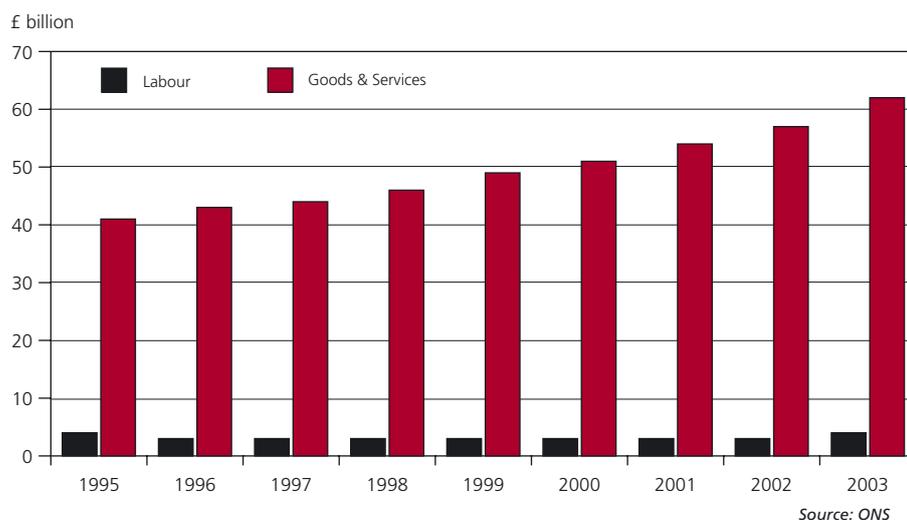
### Inputs at current prices

- 7.51** Figure 7.1 shows the trends for each input component in current prices from 1995 to 2003. (The capital consumption figure is too small to be shown on the graph.)
- 7.52** The data at current prices originate in the four UK health departments. The raw data are collated and processed by the Treasury, to feed into the government financial management system as well as to support the National Accounts. Some items which appear as expenditure in accruals accounts – provisions, for instance – do not feature as an expenditure item in the National Accounts. Such technical differences, and the need for UK figures on a calendar year basis, sometimes make it hard for any of the four health departments to recognise the precise figure for government final consumption on health care which appears in the National Accounts. This is part of the wider problem discussed in paragraph 6.29. The same recommendations therefore apply. Some of our recent work has shown the sort of mistakes and misclassifications that may go unnoticed for lack of regular dialogue between ONS and spending departments. Accuracy could be improved by recognising the need for wider ownership of the input data used in the National Accounts.

### Estimating inputs in volume terms

- 7.53** The acknowledged source of methodological guidance for estimating inputs in volume terms is the OECD Productivity Manual (see paragraph 3.13).
- 7.54** The current state of ONS's methodology is as follows:
- a) for labour inputs, expenditure on health care services is expressed in volume terms by correcting for price changes using DH's earnings index for health care workers;
  - b) for inputs of goods and services, expenditure is expressed in volume terms by correcting for price changes using a composite index. One part reflects the fact that hospital costs are largely staff-related and the other is a price index designed to reflect the purchasing pattern of hospitals, compiled regularly by DH;
  - c) In the absence of a direct measure of capital services, Health capital consumption is estimated using past capital investment trends and depreciation rates for government health care.
- 7.55** Figure 7.2 illustrates the constant price trends from 1995 to 2003. (The capital consumption figure is too small to be shown on the graph.)

Figure 7.2 Health Care Expenditure (Constant Prices)



### Improvements

- 7.56** ONS has made a number of changes to Health input data in June 2004, including reclassifying NHS Trusts into central government and reclassification of non-medical health professional education costs from Health to Education. This process has involved discussions with DH on particular issues, but we consider there is scope for more systematic ways of improving routine communications to ensure that data used for Health inputs in the National Accounts is fully informed on relevant financial and organisational changes.
- 7.57** The methods used to derive the labour component of input estimates at constant prices are very basic. Ideally, an attempt should be made to follow the OECD's recommendation of producing a weighted total of hours worked which reflects the changing skills mix. If that proves impossible, the spending figures should be deflated at a more detailed level so that the resulting index more nearly reflects changes in skill mix. We intend to work further with DH and the other health departments on these issues.
- 7.58** NHS accounts may provide a more accurate source of information about capital consumption than the ONS perpetual inventory model used now and a starting point for measuring capital services.

## Triangulation and Productivity Measurement

- 7.59** It is possible to measure the productivity of health care by dividing national accounts inputs at constant prices by outputs at constant prices, as has been done for government as a whole in previous ONS publications *Economic Trends*. This measure cannot be regarded as a total picture of productivity. We understand ONS will be publishing an article on health care productivity in July 2004, it is expected to explore other issues and data.
- 7.60** A measure of productivity from National Accounts data is bound to reflect any deficiencies in both input measures and output measures, as explored earlier in this chapter. In addition, some Health inputs are spent on prevention and public health services, education and training of health personnel, and research and development in medicine and health care. None of these should be expected to produce immediate value in terms of health care activity. Prevention aims to reduce the requirement for health care, with benefits accruing over many years. Staff training (much relevant expenditure remains within Health, despite the recent reclassification), may increase the productive capacity of the health service with a lag, as new staff take up posts. Or it may increase the quality of health care in different ways if current staff are trained in new ways of delivering care or improving outcomes. Research should improve the effectiveness of health care spending, but the link may be indirect and with a considerable time lag before there is sufficient body of evidence to change health care practice.
- 7.61** Analysis of productivity could omit current expenditure in areas not linked to current output. But it should in principle take account of past expenditure which should have accumulated human capital or capacity for better quality care and analyse whether current output is reflecting lagged past investment.
- 7.62** Analysis of productivity based on comparison of inputs and outputs should, under Principle G (see paragraph 5.35), be triangulated against other evidence. Movements in certain ratios are likely to be associated with increases or falls in productivity. Fruitful areas of investigation include:
- a) examining trends in the average length of stay in hospital: a fall could indicate an increase in productivity;
  - b) observing the changing skills mix of health care staff: an increase in the overall level of skills gives a legitimate expectation of increased quality of output, for which evidence can be sought; other shifts in skill mix may indicate more effective use of support staff to increase the productive time of those with more specialist skills;
  - c) examining the migration of treatments from expensive settings (eg an acute hospital) to cheaper ones (eg GP's surgery);
  - d) examining changes in the technology of treating particular diagnoses, particularly in areas where such developments move the output from one category to another, eg the substitution of drugs for invasive surgery; and

- e) levels of use of scarce resources, such as intensive care beds, ambulances or operating theatres; a very high rate of use may be as inefficient as a very low rate, as time is wasted in hunting for an alternative, and/or patients have adverse outcomes.

### Summary of Work in Progress and Future Directions

**7.63** Our main concerns are that current measures of health care used in the National Accounts:

- a) do not include quality measures, which are an essential part of output (as distinct from activity);
- b) assume that health care activity growth in England is a proxy for all UK activity, which may not be correct; and
- c) rely on the General Household Survey to estimate outputs of general practitioners and practice staff.

**7.64** During the remaining period of our review, we will undertake the following work.

- a) We will work with DH on sensitivity testing, using any sample or local data which may be available, to consider which health care activities not currently measured nationally should be particular priorities for better measures.
- b) We will work with ONS experts in the General Household Survey to consider what the error rates may be in the estimated GP consultation rates and GDP estimates based on this source. If this confirms a need for improvement, we will review with DH whether better methods can be found to measure changes in services provided by GPs, taking account of the new GP contract and other structural changes. One possibility may be the use of a long-term database of information about a subset of GP practices such as the Royal College of General Practitioners spotter practices or General Practice Research Database, though the practices involved may not be a representative sample.
- c) We will consider further how to make use of an approach which uses common cost weights for treatments which are substitutes, especially if there are differences in quality of outcome for the patient and if all costs and benefits of a complete course of treatment can be identified. This might be part of a triangulation approach.
- d) We will continue discussions with the York/NIESR research team who are advising DH on measuring NHS outputs and productivity, so we can take account of their work (which continues until late 2005) when we comment in our final report on the most promising directions for incorporating quality measures into Health outputs as used in the National Accounts.
- e) We will work with ONS, DH and the health departments in Wales, Scotland and Northern Ireland to see how far relevant data from their respective health systems can be incorporated into the UK National Accounts, using the techniques adopted for England in 2004 or adapting them as appropriate, given differences in systems, under the principles set out in this report.

f) We will work with ONS and DH to improve techniques to estimate labour and capital components of input estimates at constant prices, including more accurate ways of reflecting changes in skill mix and relative earnings.

**7.65** For the longer term, we see benefits in further work to explore new measures of NHS work which combine separate activities into ‘care pathways’, to give the clearest picture of value added by the NHS. We consider it important that developments in NHS information systems, while intended primarily to support patient care, should also support better future analysis of health care outputs.

# 8 Education: A Progress Report

## Introduction

- 8.1** Expenditure on schools in the United Kingdom involves both central and local government. This includes expenditure on nursery, primary and secondary education in schools, which are the responsibility of local education authorities. In 2000, Education accounted for 17.1 per cent of general government expenditure. It is the second biggest area of expenditure, after Health, and made up 3.2 per cent of GDP in 2000.
- 8.2** Education output is currently measured by the number of pupils taught, plus a fixed assumed quality adjustment. Increased expenditure on education in recent years has been aimed at improving the quality of education provided, to improve literacy and numeracy. It is not clear whether such improvements have been adequately measured in the fixed quality adjustment.
- 8.3** The Department for Education and Skills (DfES) is responsible for education in England. The department states that its aim is ‘to help build a competitive economy and inclusive society by creating opportunities for everyone to develop their learning; releasing potential in people to make the most of themselves; and achieving excellence in standards of education and levels of skills.’
- 8.4** The department’s objectives are to ‘give children an excellent start in education so they have a better foundation for future learning and enable all young people to develop and to equip themselves with the skills, knowledge and personal qualities needed for life and work.’
- 8.5** The 2002 Public Service Agreement (PSA) targets relating to schools are:

Objective I: sustain improvements in primary education

Raise standards in English and maths so that:

- by 2004, 85 per cent of 11 year olds achieve Level 4 or above and 35 per cent achieve Level 5 or above, with this level of performance sustained to 2006; and
- by 2006, the number of schools in which fewer than 65 per cent of pupils achieve Level 4 or above is significantly reduced.

Objective II: transform secondary education

Raise standards in English, maths, ICT and science in secondary education so that:

- by 2004, 75 per cent of 14 year olds achieve Level 5 or above in English, maths and ICT (70 per cent in science) nationally, and by 2007, 85 per cent (80 per cent in science);
- by 2007, the number of schools where fewer than 60 per cent of 14 year olds achieve Level 5 or above is significantly reduced; and
- by 2007, 90 per cent of pupils reach Level 4 in English and maths by the age 12.

Objective III: pupil inclusion

- by 2004, reduce school truancies by ten per cent compared to 2002, sustain the new lower level, and improve overall attendance levels thereafter; and
- enhance the take-up of sporting opportunities by 5-16 year olds by increasing the percentage of school children who spend a minimum of two hours each week on high quality physical education and school sport within and beyond the curriculum from 25 per cent in 2002 to 75 per cent by 2006. Joint target with Department of Culture, Media and Sport.

Objective IV: raise attainment at 14-19

Raise standards in schools and colleges so that:

- between 2002 and 2006, the proportion of those aged 16 who get qualifications equivalent to five GCSEs at grades A\* to C rises by two percentage points each year on average and in all schools at least 20 per cent of pupils achieve this standard by 2004 rising to 25 per cent by 2006; and
- the proportion of 19 year olds who achieve this standard rises by three percentage points between 2002 and 2004, with a further increase of three percentage points by 2006.

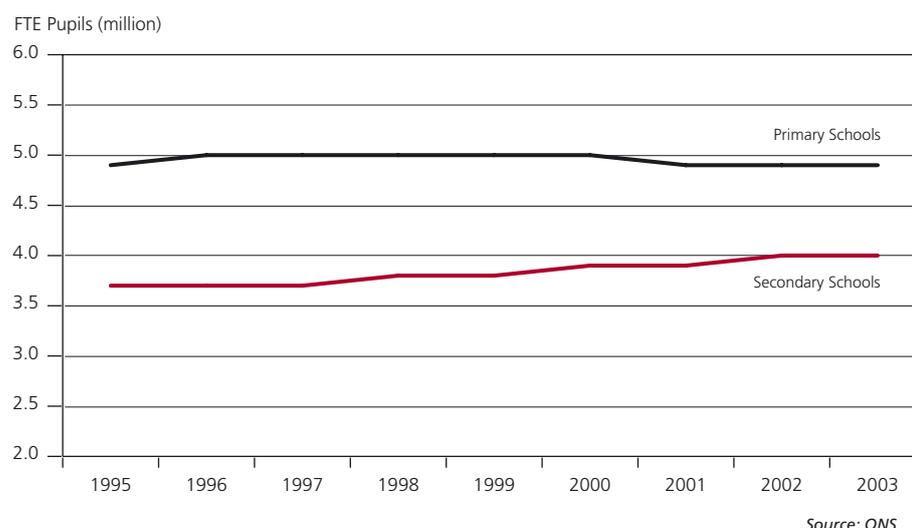
- 8.6** Paragraphs 8.33-8.49 discuss ways of developing an output measure which would make use of examination attainment in some of the areas covered by the PSA targets.

## How UK Education Output is Currently Measured

### Schools

- 8.7** Direct measurement of Education output was introduced into the National Accounts in 1998, backdated to 1986. The measure of the volume of services delivered was full-time equivalent (FTE) pupil numbers weighted by cost. A quality adjustment of 0.25 per cent was introduced at the same time (see paragraph 8.11).
- 8.8** Measures of non-market general government output for Education relate to maintained schools in England, Wales, Scotland and Northern Ireland. Non-market services are those provided free at the point of delivery. In the current United Kingdom education system this includes maintained nursery schools and nursery classes through to sixth form education at maintained schools. It does not include further education or higher education which are provided by bodies outside the government sector of the National Accounts.
- 8.9** The number of pupils is used to measure the volume of education. The FTE numbers of pupils in nursery, primary, secondary and special schools in the United Kingdom are collected annually. Pupil numbers in each of the four types of school are weighted by expenditure for that type of school in the base year (currently 2000). Expenditure in primary and secondary schools covers 90 per cent of total expenditure in the non-market education sector.
- 8.10** Overall pupil numbers are fairly stable but there are underlying trends. From 1995 to 2001, pupil numbers increased each year, by 3.8 per cent over the six year period. From 2001 to 2003, the numbers fell by 0.3 per cent over the two years. This is due to a decline in primary school numbers; secondary school numbers are still rising as shown in Figure 8.1.

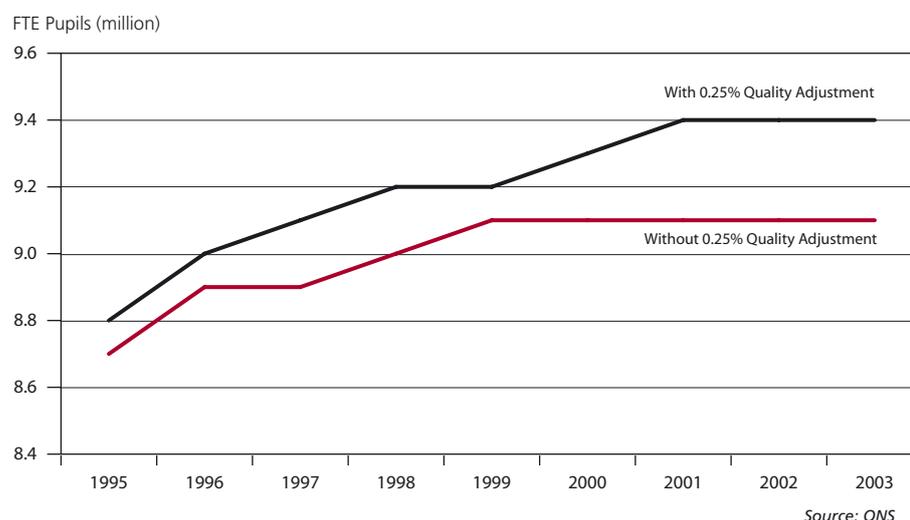
Figure 8.1 **Number of FTE Pupils in UK schools**



**8.11** The method of quality adjustment for Education introduced in 1998 was based on changes in exam results. The justification was that the quality of educational services delivered could be proxied by exam success. A quality uplift factor of 0.25 per cent is used for UK primary and secondary schools each year. The outcome indicator used was the average points score of pupils aged 16 in England using the standard GCSE scoring system. The change recorded over a period of four to five years was averaged to get an estimate of annual change in points score. There has been no subsequent recalculation of this adjustment to verify whether the same improvement in points score was achieved by subsequent cohorts of pupils.

**8.12** Figure 8.2 shows the trends of General Government Education output from 1995 to 2003. The top line shows the output trend with the 0.25 per cent quality adjustment followed by the output trend with no quality adjustment. Both time-series are cost weighted.

Figure 8.2 **Total UK Education Output**



### Education of non-medical health professionals

**8.13** Government expenditure in the National Accounts is classified by purpose, rather than by the responsible government department. From June 2004, a change has been made to the National Accounts to include the Department of Health's spending on non-medical health professional education in the Education category of government expenditure rather than Health. The expenditure goes from NHS bodies to universities to purchase places on diploma or degree courses in nursing and allied health professional training. Many of these are three-year courses taken by students who are not employed by the NHS but intend to make a career in health care. Some students may be current NHS employees taking further qualifications. (University education in medicine and dentistry is funded differently and is outside general government expenditure.)

- 8.14** ONS introduced in June 2004 a direct measure of Education output for non-medical health professional education, taken from a series for the numbers of health students being trained. When the expenditure was included previously in the Health category of the accounts, there was no separate output measure. Health output reflected only the cost weighted activity index described in paragraphs 7.9-7.10 with no measure of output from training activities.
- 8.15** NHS expenditure on non-medical health professional education in England is used as a proxy for UK expenditure: expenditure per head of population in England is grossed up by the population of England, Wales, Scotland and Northern Ireland.

### International Context

- 8.16** The Education classification contains a wide diversity of institutional units from schools and universities to vocational institutes and private driving schools. These are provided by various levels of government, non-profit institutions and non-financial corporations.
- 8.17** The *Eurostat Handbook on Prices and Volume Measures in National Accounts* considers all education services to be individual services where the output measure should focus on individual benefit for pupils. It views the inputs of the Education sector as producing school lessons, the activity. The school lessons are, in turn, taught to pupils, where the number of pupils taught is seen as the output of Education. This output together with external factors result in achievements in examination results - the final outcome of the education system.

‘Education is an “individual service”, delivered to pupils by educational establishments, which themselves use the inputs of staff time, capital consumption, and intermediate consumption ... Teaching is generally undertaken in groups, which can range from very small groups with close coaching to large seminar or lecture groups of over 100 pupils. But the output measure should focus on individual benefit for pupils and not be based on the group level.’ (Eurostat Handbook, section 4.12.)

- 8.18** The unit of output (the preferred Eurostat approach) is not always easy to define. Output can be seen as an actual delivery from the producer to the consumer. The Eurostat Handbook warns that measures of social factors such as the level of educational attainment may be affected by influences other than the quantity or quality of publicly funded education such as parental involvement or private tuition (see paragraph 5.5). The Eurostat Handbook defines Education output as:

‘... the quantity of teaching received by the student, adjusted to allow for the qualities of the services provided, for each type of education.’

‘the quantity of teaching received by students can be measured by the number of hours they spent at being taught. This measure is referred to as the number of “pupil-hours”. Where this measure is not available, the simple number of pupils can be an alternative.’

As noted in paragraph 5.4, we are not altogether persuaded by the Eurostat reasoning. There looks to be some potential confusion between activities and genuine outputs. At the least, this could stand further examination.

Appendix C provides further information on the international context.

**8.19** All EU countries, except Denmark, are committed to measuring the output of Education via the direct volume measure by 2006. We discussed Education output and input measures during visits to the Statistical Offices of Italy, the Netherlands, Finland, Norway and Sweden. In line with international guidelines, the volume measure used in Italy, the Netherlands and Australia is, like the United Kingdom, the number of pupils, weighted by cost.

**8.20** Eurostat guidance recognises a number of possible indicators based on ‘outcome’ and ‘output’ measures of quality. No single recommended method to adjust for quality is made, as different education systems can lead to different models for taking quality properly into account. As a result, quality is adjusted in various ways in the countries that use output measures, as shown in Table 8.1.

Table 8.1 **Countries using or planning to use direct output measure**

<i>Countries currently using direct Education output measure</i>		
<b>Country</b>	<b>Activity measure</b>	<b>Quality measure</b>
UK	Numbers of pupils, weighted by costs	Constant adjustment based on exam success
Italy	Numbers of pupils, weighted by costs	Class size – smaller is better quality
Netherlands	Numbers of pupils, weighted by costs	Moving on to next academic level
Australia	Numbers of pupils, weighted by costs	None
<i>Countries planning to adopt direct Education output measure</i>		
<b>Country</b>	<b>Activity measure</b>	<b>Quality measure</b>
Finland	At planning stage; likely to be number of pupils (to be published from 2006)	None
Sweden	Pupil hours, weighted by cost of each type	Exam failures do not contribute to output
Norway	Work is at planning stage	
Germany	Pupil hours of education (to be published from 2005)	None, researching methods

Appendix D, paragraphs D21-D50, provide further information on the other countries.

## Critique of Current Education Output Methods

- 8.21** The UK methodology for estimating output of schools could be improved in a number of ways, in line with the criteria in paragraph 6.3.

### Capturing all outputs

- 8.22** There is complete coverage of the number of pupils in UK nursery, primary and secondary schools. The stratification by category corresponds to that required in the Eurostat Handbook (Eurostat Handbook, section 4.12). The ideal indicator of output is the number of hours spent in education by pupils, with an adjustment for quality. Where this measure is not available the number of pupils is seen as an acceptable alternative, so the United Kingdom use the FTE number of pupils. However, it may be possible to approximate number of hours using information on length of school day and school absence levels by type of school. It is worth reviewing whether appropriate statistics are available and examining what difference they would make to the output measure. This would allow us to check whether the amount of hours being taught is sufficiently stable over time, as required by the Eurostat Handbook (Eurostat Handbook, section 4.12).
- 8.23** The inclusion of non-medical health professionals' university education in 2004 suggests there would be value in an investigation into whether any other areas of government expenditure should be re-classified into Education.

### Improving measures of output to capture changes in quality

- 8.24** While it is commendable in principle that the United Kingdom applies a quality adjustment to the volume measure of pupil numbers, it is less satisfactory that the standard 0.25 per cent increase has been applied in each year since 1998 without re-examining the methodology to see if it still reflects more recent performance of the education system. The next section describes ways forward to make use of up to date measures of pupil progress.
- 8.25** Further thought needs to be given to whether a rolling update, which reflects results from several recent years, might be preferable to a measure which relies entirely on the latest year's examination results. That could be important if there are changes in examination systems or doubts about whether standards are entirely consistent between years (see paragraph 5.47).
- 8.26** In principle, it seems better to recognise health professional education as a distinct output rather than subsume it within Health (or school pupil numbers, if no separate output measure had been introduced after re-classification). However, it may be possible to improve the measure introduced in 2004. The new measure is a simple count of student numbers. It does not distinguish between courses of different duration, cost, or potential value to the NHS. It has no quality adjustment. Measures like the proportion of students completing the course and taking up the relevant professional registration could be regarded as quality measures.

### Improving UK coverage and timeliness of information

- 8.27** The current volume measure uses data on pupil numbers from all parts of the UK. However, the current quality adjustment was based on examination results from England only. Ideally a new quality measure should be based on information from England, Wales, Scotland and Northern Ireland. There are important differences in examination systems and other aspects of education so careful exploration will be needed, with the respective devolved administrations, to choose an appropriate approach.
- 8.28** We expect DfES to continue working with ONS to ensure timely transfer of appropriate data for the purposes of National Accounts.
- 8.29** The new non-medical health professional output measure ought to have UK coverage.

### Ways Forward on Output Methods

- 8.30** Our main concern is to improve the way of measuring quality as part of school education output for England, and then move on to adopt similar approaches for the other UK countries, recognising the different arrangements and different information systems for education in each country.
- 8.31** We expect pupil hours or the number of pupils in schools to continue as the volume measure of Education activity. But changing birth rates and increasing participation rates in nursery and post-16 education both affect total numbers, and some weighting by cost of school or length of school day is important to reflect the different balance within overall pupil numbers.
- 8.32** Various approaches to measuring quality can be considered in line with international guidance. Quality can be measured through outcome-based indicators including examination data, perhaps with a 'value-added' approach, where pupil attainment is measured before and after a period of schooling, or from standardised assessment tests. Quality adjustment could also be measured through measures of output quality such as provided by school inspections. Pupil/teacher ratios can be proposed as indicators of quality where this is based on a sound analysis of the relationship between the ratio and the quality of classroom teaching. We consider each of these approaches in this section, based on work which so far covers only England. The eventual choice of methods must be based on a rigorous statistical and expert analysis of the alternatives available.

**Box 8.1 Literature on Measurement of Outputs of Education**

Vignoles, A *et al* (2000), in *The relationship between resource allocation and pupil attainment: a review*, assessed many of the papers on the linkage between expenditure, quality and pupil attainment. As a whole, the research found that some measurable school inputs do matter. These include class size, teacher experience and teacher salaries. However, the magnitude of the effects found was seen to be small.

Education output can be measured through the rate of return on human capital acquired through education. O'Mahony and Stevens (2003) projected future earnings for men and women in the United Kingdom and the United States with various qualifications using a cross sectional regression method. From 1994 to 2001, annual earnings rose because of increases in productivity by 2.1 per cent in the United Kingdom and 1.5 per cent in the United States. The estimates suggest 0.5 per cent in the United Kingdom and 0.2 per cent in the United States was due to education.

Ahlroth *et al* (1997) and Ervil A *et al* (2003) have examined the lifetime earnings generated by the education system in Sweden and Norway. Both found that there were marked differences between the estimated figures and the conventional approach used within the National Accounts. If implemented, the new estimates would have changed the picture of the overall performance of the respective economies. (See Appendix D39-D40 for further details.)

Various research papers on performance of the educational sector at the micro level using Key Stage (KS) results are helpful in development of a new output measure.

Goldstein *et al* (1999, 2000 & 2003) use a multi-level model to show that schools differ in several ways, including by curriculum subject and prior attainment of pupil. Goldstein argues that measurements of schools' performance should take account of students' achievement prior to entry, as some schools are highly selective – either deliberately or inadvertently.

The NAO (2003) used multi-level modelling to assess the performance difference between secondary schools after adjusting for prior achievement and some external influences.

Blatchford, P, Goldstein, H, Martin, C and Browne W, (2002) in 'A Study of Class Size Effects in English School Reception Year Classes' *British Education Research Journal*, Vol 28 followed two large cohorts of children over the first three years of school. The authors used a series of multi-level models to explore the correlation between achievement and class size. The results showed a clear effect of class size on children's academic achievement in their reception year. There was a larger gain from being in a small class for lower baseline achievers compared to the other ability groups.

### Quality of education and pupil attainment

- 8.33** A possible method for measuring the quality of school output is to consider the achievements of pupils in the statutory Key Stage tests. We can do this by either looking at the absolute results of the pupils or by looking at the rate of progress of the pupils. The quality measure of Education output is therefore based on the performance of pupils in the system. In the view of DfES, this is the best indicator of quality. The use of exam results to measure quality also fits with existing DfES PSA targets.
- 8.34** In the current compulsory education system in England, pupils are required to sit statutory tests at ages 7, 11 and 14 and GCSE exams at age 16. The tests are called Key Stage 1, Key Stage 2, Key Stage 3 and Key Stage 4 (or GCSE) respectively. The tests aim to assess the knowledge and skills gained by pupils over the period of the key stage. Table 8.2 provides details of the key stages.

Table 8.2 **Comparison of key stages and schooling years**

Key Stage	School Year	Age of pupil at start of year	Number of years between tests
1	2	6	(3 years since entry)
2	6	10	4
3	9	13	3
4 (GCSE)	11	15	2

- 8.35** There are a number of permutations on how output can be calculated using examination results. We can either base the measure on GCSE results alone or as a combination of all the key stage outcomes. Either method can be calculated using either raw attainment results or pupil progress figures. The various approaches to measure the output of the system are listed below.
- Use GCSE results, with either percentage of pupils achieving a threshold measure or average point score results.
  - Use Key Stage 1,2,3 and GCSE results as indicators of the output of different stages of schooling.
  - Use Key Stage 2 performance as an output indicator for primary education and value-added performance from Key Stage 2 to GCSE as an indicator for secondary education.
  - Use a pupil progress measure to show the level of improvement between each of the key stages and GCSE results, and use actual Key Stage 1 performance as an indicator for Key Stage 1.
- 8.36** An index of quality could be produced from the annual results; key stage results would be weighted using the number of pupils within each key stage in order to produce an aggregate index.
- 8.37** DfES collect a large amount of information annually on qualifications achieved. However, using this information to obtain an annual measure of quality for National Accounts is not straightforward.

### GCSE results

- 8.38** GCSEs provide a measure of the final outcome of compulsory education. This could be based either on the percentage of pupils achieving a threshold number of passes of certain grades, as required in the PSA targets, or on the number of points achieved. (For each GCSE, eight points are allocated for an A\* grade, declining to one point for a G.) Both measures have risen in each year to 2003. The proportion of pupils achieving the threshold of five or more GCSEs at A\* to C has risen steadily each year, from 43.8 per cent in 1995 to 52.9 per cent in 2003.
- 8.39** The GCSE result covers eleven years of education. The rationale for using this measure would be that it does not need to assume or estimate the stage at which improvements have occurred. Good GCSE results may not be due to the improvement from Key Stage 3 to GCSE but to an improvement made at an earlier key stage. Another advantage is the simplicity of the single measure.
- 8.40** A disadvantage is that this measure is less representative of recent changes in primary and nursery education. It would take many years before productivity changes in the earliest key stages show through at GCSE and it would not be possible to attribute these changes to the correct time-period.

### Regression analysis

- 8.41** DfES are researching the use of regression techniques as a means of measuring productivity in primary and secondary schools. This is an alternative technique for measuring pupil progress, controlling for prior attainment. Here the outcome at a particular key stage would be regressed on prior attainment at a previous key stage. The coefficient on the prior attainment variable would reflect the actual effect that a given change in, say, Key Stage 3 had on GCSE results. Over time, this relationship would shift with changes in overall school productivity. The change could be measured by the inclusion of time dummy variables. This approach could be applied either to pupil level or school data.
- 8.42** This is a new method of measuring productivity and its reliability would need to be carefully evaluated. DfES do not currently recommend it as a basis for a National Accounts measure.

### Transition matrices

- 8.43** DfES also measures pupil progress using transition matrices (Box 8.2). They have devised a way that combines pupil progress between Key Stages 1-2, 2-3 and 3-4 with Key Stage 1 raw attainment; without a baseline test at school entry, Key Stage 1 results are the best measure available of progress up to Year 2.

**Box 8.2 Preliminary Results of the Transition Matrices**

Pupils are tested in a series of four key stage tests up to the age of 16. Pupil progress between successive key stages is described by *transition matrices*.

Progression of Pupils from 2000 KS2 to 2003 KS3

Table B8.2 **Progression of pupils from 2000 KS2 to 2003 KS3**

		2003 KS3 Average Levels							Total
		B3	3	4	5	6	7	8	
2000 KS2 Average Levels	B3	32	57	10	1	0	0	0	100
	3	2	30	53	15	1	0	0	100
	4	0	2	16	48	32	2	0	100
	5	0	0	0	7	50	42	0	100
	6	0	0	0	0	2	97	2	100

The Transition Matrix gives the percentage of pupils that move nationally from each average level at Key Stage 2 in 2000 to each of the levels at Key Stage 3 in 2003. Here it can be seen that in 2003, 48 per cent of pupils who achieved the expected level (Level 4) at Key Stage 2 in 2000 went on to achieve the average expected level (Level 5) in Key Stage 3 three years later and 32 per cent went on to achieve a Level 6.

Each year a Transition Matrix can be created for pupil changes from Key Stage 1 to Key Stage 2, Key Stage 2 to Key Stage 3, and Key Stage 3 to GCSE.

- 8.44** DfES have also devised a method of converting key stage results into GCSE equivalents, by ‘moving’ pupils at any key stage through the appropriate set of transition matrices, thus generating hypothetical future GCSE outcomes for younger children.
- 8.45** The method then compares what was actually achieved in a particular year to what would be expected, assuming progress remained constant. The output change of Key Stages 2, 3, and 4 can be calculated in this way. The total output change in any year is the weighted sum of these output changes along with Key Stage 1 changes.
- 8.46** The method uses the current year’s results but could be extended to include other years to be more representative and to smooth out the effects of any erratic results.
- 8.47** When pupils achieve an improvement at one stage, they do well to sustain this improvement through to the next stage. Improvement achieved at one key stage tends to reduce pupil progress at the next level. Therefore, while there may be dramatic improvements in terms of say Key Stage 2 results over a few years and although Key Stage 3 results will continue to improve as this cohort progresses into Key Stage 3, the pupil progress from Key Stage 2 to Key Stage 3 (in relation to the expected standards) will nearly always decrease. This means that an increase in pupil progress may be followed by a subsequent decrease even though overall attainment might be continuing to rise. One way to counter this may be to smooth the achievements over several years.

**8.48** The tests and exams suggested are based on pupils of compulsory school age. The method devised does not include post-16 education in secondary schools. This is a weakness, given that extended participation at nursery and post-16, and increased numbers of pupils passing A levels at 18 and entering university, are major government priorities and funding has been allocated accordingly. Unless A level results are included in some way in a quality adjustment, the National Accounts would not take account of educational achievement after 16. Extra spending on relatively expensive tuition for over 16 year olds could depress the productivity of the school system, if there is no corresponding quality adjusted output.

#### Pupil attainment assessment

**8.49** There are a number of issues surrounding the use of examination results.

- a) Examination systems can change over time. In the future, new tests or other forms of assessment could replace the key stage tests and GCSEs.
- b) Examination standards may not be consistent over time. Improving results may reflect 'teaching to the exam' rather than wider educational attainment.
- c) Examination results are influenced by a range of non-education factors. Ideally the measurement should only include pupil attainment that can be attributed to schools. Research is being conducted at DfES into methods for taking out non-school factors. However, for national results, the assumption could be made that non-school effects are constant between one year and the next and so would not affect national trends.
- d) Examination results do not measure everything schools do. Key stage test results relate only to English, maths and science and do not necessarily correlate with all round success in other subjects or preparation for adult life. Personal and social skills such as self-reliance, team working, communication skills and inter-cultural understanding are also an important output of schooling, but are hard to measure and value.

#### Valuing education

**8.50** DfES have also been researching the valuation of educational outputs in terms of discounted lifetime earnings. This assumes that education increases future labour market earnings, which gives a way of measuring the value of the education provided. GCSE results can be valued using earnings data in the Labour Force Survey. In order to match the survey data, GCSE results have to be placed in bands, which range from no GCSEs to at least five at A\*-C. The earning profiles for each of these groups can be computed.

**8.51** Another dimension of education value is that it provides child care services. Estimated charges to parents for child care are used as a proxy for the value of child care services to children in primary education aged from five to seven. The charges of after-school clubs are used to estimate the child care benefits accruing in education between the ages of 8 and 14. Preliminary calculations by DfES estimate childcare benefits ranging between £17.5bn and £18.3bn a year in the period 1995-2003.

### Quality of schools and teaching based on Ofsted reports

- 8.52** Quality of education could be measured through school inspections, using the quality scores awarded to schools in England in inspections by the Office for Standards in Education (Ofsted – see Box 8.3).

#### Box 8.3 Role of Ofsted

The Office for Standards in Education (Ofsted) was set up on 1 September 1992. Ofsted is a non-ministerial government department headed by Her Majesty's Chief Inspector of Schools in England (HMCI). Ofsted is independent of other government departments, although it works closely with them. This independence enables educational matters to be inspected and reported impartially. Ofsted regularly inspects all the 24,000 schools in England, which are wholly or mainly state-funded. Ofsted also reviews local education authorities (LEAs), inspects teacher training courses, child care, the private, voluntary and independent nursery sectors. It also inspects independent schools (including independent special schools) and service children's education, and reports on LEA-funded youth services.

- 8.53** Ofsted reports grade the school and quality of education provided, based on recorded evidence which must be of sufficient range and quality to justify the judgement. Reports use a seven point scale (see Table 8.3), though this is soon to be changed to a four point scale. As well as individual school reports, statistics are produced annually summarising inspection results for different types of school: nursery education, primary schools, secondary schools, post-compulsory education and special schools.

Table 8.3 **Ofsted seven-point judgement scale**

Quality descriptor and alternatives	Grade	Implications
<b>Excellent:</b> exceptional; outstanding; first-rate; very effective; very rapid	1	Worth disseminating beyond the school
<b>Very Good:</b> well above average; highly effective; rapid	2	Worth sharing within the school
<b>Good:</b> above average; effective	3	Worth reinforcing and developing
<b>Satisfactory:</b> average; acceptable; sound; typical	4	Adequate, but scope for improvement
<b>Unsatisfactory:</b> below average; inadequate; slow; ineffective	5	Needs attention
<b>Poor:</b> well below average; very ineffective; very slow	6	Needs urgent action
<b>Very Poor:</b> extremely ineffective; extremely slow	7	Immediate radical change needed

Source: Ofsted

- 8.54** The scores awarded to individual schools in the inspection reports could be summarised to measure changes in the quality of education. For example, the percentage of schools graded as good and above for teacher quality could be used as an annual measure of quality.

**8.55** But using Ofsted inspection reports as a measure of the quality of education raises a number of difficulties:

- a) Ofsted continually reviews and changes its inspection procedures. There was a major change in grading in 1997 (from a five-point to a seven-point scale) and the next change is planned for September 2005 (back to a four-point scale). This makes long term comparisons difficult.
- b) Ofsted scores are categorical variables: the difference between score three (good) and four (satisfactory) may not be the same as between score four and score five (unsatisfactory), and any crude aggregation of points scores could be invalid.
- c) Schools prepare thoroughly for Ofsted inspection and there may be an artificial climate, subjective professional judgements and no certainty that the criteria for teaching and school quality would be consistent across schools and time.
- d) Schools are not selected randomly for inspection. All schools are inspected at least once every six years, but weaker schools are selected for inspection more frequently, biasing the sample of results each year.

**8.56** There are, though, a number of advantages which justify further exploration of this source of information about the quality of Education output. Ofsted have an unrivalled depth and breadth of knowledge about the quality of schools. Inspections cover all aspects of education, not just examination scores. When making assessments they take account of particular challenges facing individual schools when making assessments, and consider schools' contribution to improving and promoting social inclusion.

**8.57** Ofsted advised us that, although frameworks for inspection have evolved, the method of judging teaching quality had been one of the more stable elements. There are checks and balances to ensure consistency in inspection standards, with inspectors working to a common framework and themselves subject to inspection and scrutiny.

**8.58** Ofsted ensure that its annual summary of inspection results is nationally representative by using appropriate weights for the numbers of different types of schools inspected.

#### Quality of resources based on adult/pupil ratio

**8.59** An alternative approach for measuring quality would be to use class size, or pupil/teacher ratio, on the assumption that the smaller the pupil/teacher ratio the better the quality of learning. But clear evidence of these relationships would have to be established.

**8.60** Class sizes in infant schools (for five, six and seven year olds) were reduced to a statutory maximum of 30 from September 2001 as part of the drive to raise standards in schools. There are no plans to require further reduction in class sizes, which have remained fairly stable over recent years. There may be advantages in having smaller classes in some circumstances but the case is not proven that reducing class size alone would improve quality.

**8.61** The number of teachers and the pupil/teacher ratio have been fairly constant over recent years. However, schools in recent years have taken on a greater number of classroom support staff. One of the objectives is to release teachers' time from tasks that support staff can undertake (eg taking registers) and so increase productive time for trained teachers. Support staff can also directly help pupil learning. So the overall increase in school staff should arguably have improved quality of education. It would, though, be wrong to presume that outputs have changed by measuring change in inputs. Examination of changes in labour productivity as the skills mix in schools changes would be useful in the triangulation approach to measuring productivity. It will also be important to take proper account of the pay of classroom support staff in deflating educational expenditure.

### Summary

**8.62** None of the methods outlined can yet provide an exact or complete measure of the quality of Education output delivered to individual pupils, as required for the National Accounts. We intend to work further, with DfES, on a number of approaches. Our conclusion may be, after further work, that more than one approach is viable and that a weighted measure could be used. If so, we would need to consider carefully what their relative weights should be.

**8.63** Both Ofsted reports and examination results have shown improvements over the past ten years. The current tests for Key Stages 1, 2 and 3 and GCSEs have been in operation for a number of years and provide good trend data. There is, so far, less data for pupil-level value added results.

**8.64** A concern for any proposed measure is the likelihood of future major change in the system or source data, and also the sensitivity of the measure to minor changes, past or future. The Ofsted grades will change in 2005. Examinations at 16 may change as a result of the current 14-19 education review. That may also change examinations at 17 and 18, which would affect measurement of value added for this age group.

**8.65** Where change is unavoidable, it is important to plan in advance for measurement during transition, including sensitivity testing and the use of dual measures where possible. An attraction of pursuing two (or more) measures is that, in such a case, the other measure(s) can be used for calibration and perhaps for dampening the effects of transition – if they do not change simultaneously.

**8.66** Quality adjustments also need to be made to the weighted pupil numbers for Scotland, Wales and Northern Ireland. Although the broad methodology ought to be consistent across the United Kingdom, there are differences between educational systems, including the timing and nature of examinations or key stage tests, and these may affect the approach to be taken. We intend to discuss the issues further with the devolved administrations during the remainder of the review, to see how we can improve on an assumption that quality change in the English educational system is an accurate proxy for change in the other countries of the United Kingdom.

**8.67** We note with interest the possibility of valuing Education outputs from future earnings and value of child care provision. Both approaches are consistent with the comment in paragraph 5.54 that government outputs may increase in value in line with wider economic growth, eg with educational attainment increasingly valuable to individuals in a knowledge-based economy. We will consider whether these wider economic valuations could form part of the measure of Education output used in the National Accounts; even if not directly useable in the National Accounts, they would form part of a useful wider data set for triangulation.

## Inputs and Deflators

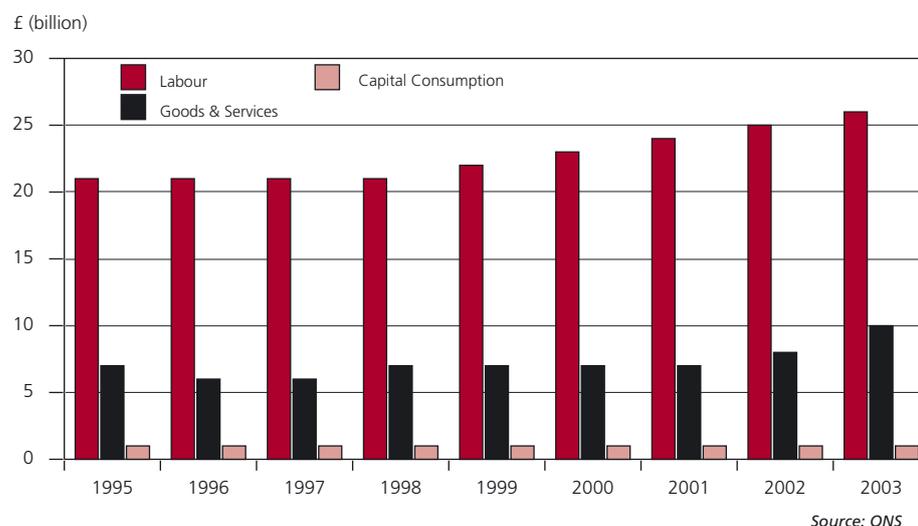
### Current methods used

**8.68** The review is also examining the appropriate measurement of Education inputs. It is looking at the distinction between current and capital expenditure and at current expenditure which is designed to build capability over time rather than support current activities, such as staff training and development. It is also considering price indices to ensure that the most appropriate adjustment is made to expenditure when calculating constant prices.

**8.69** The three components that make up total input expenditure are labour, intermediate consumption of goods and services and capital consumption. These are summed to give Total Final Consumption Expenditure in current prices. To achieve a constant price figure, the money value of these components must be adjusted (using appropriate price indices) so that the effect of inflation is taken out leaving the real changes in the volume of the inputs. The deflation process is carried out separately for each component before the inputs are re-aggregated.

**8.70** Figure 8.3 shows the trends for each input component in constant prices from 1995 to 2003. In 2002, labour accounted for around 73 per cent of Education inputs, goods and services for 24 per cent and capital consumption for three per cent.

Figure 8.3 **Education Expenditure (Constant Prices)**



- 8.71** Ninety-five per cent of Education labour costs are in local government. Expenditure data is reported annually by the Office of the Deputy Prime Minister (ODPM) for England and by the devolved administrations. UK local government education labour costs are deflated to constant prices using a local authority education pay index supplied by ODPM. This is based on data from England only, using headline pay settlements for teachers and for local government staff, together with National Insurance and pension contributions. It is assumed that pay increases are similar for the devolved administrations. The remaining five per cent of education labour costs are in central government. Data are taken from the Treasury Public Expenditure Statistics (PES) database and is deflated using the public sector average earnings index.
- 8.72** Expenditure on goods and services is the second biggest input component for Education. In 2003, over 80 per cent of total expenditure on goods and services was spent by Local Government with the remaining 20 per cent spent by Central Government.
- 8.73** Detailed data on Central Government Education procurement are unavailable and so a single net procurement figure is used. This is deflated to constant prices using a composite price index of 17 Producer Price Indices (PPI) (for goods) and the Retail Prices Index excluding mortgage interest payments (RPIX) (for services).
- 8.74** Local Government Education capital consumption – about 98 per cent of the total – is deflated using a composite deflator. This composite deflator is constructed using the public sector deflator for buildings, plant and machinery deflator for private education, PPI for vehicles and construction. These deflators are weighted by the capital expenditure of land, the plant, vehicles and construction for education. Central Government Education capital consumption is deflated using an implied deflator for total Central Government capital consumption.

#### Critique of current methods

- 8.75** Given that labour accounts for over 70 per cent of total Education expenditure, it is important to use the most appropriate pay deflator. The ODPM pay index does not use data from Scotland and Northern Ireland, which are not covered by pay negotiations for teachers and local government staff in England and Wales, and uses headline pay settlements where it would be preferable to deflate by changes in average earnings.
- 8.76** The preferred index for deflating labour costs would take into account employer labour costs, pension contributions and National Insurance, as well as earnings. Ideally, the methodology should also take into account the skill of the employee. In Education, there are four main occupational classifications:
- teachers;
  - civil servants;
  - local government administrative and clerical staff; and
  - other staff.

- 8.77** In recent years, the implied deflator for Education goods and services has remained relatively flat when compared with the RPIX. This could indicate that the deflation process is not adequately adjusting for price movements, resulting in the volume of these goods and services being over estimated.
- 8.78** The ideal method for deflating goods and services is to apply specific price indices on a product-by-product basis, or to use a specific Education procurement index. Work to facilitate this would be deemed highly desirable. An improved Education capital consumption deflator would also be desirable if a source could be found.

### Triangulation and Productivity Measurement

- 8.79** Triangulation – making use of evidence on Education quality and productivity from several sources as outlined in paragraphs 5.33-5.36 – will be an important way to interpret and calibrate data collected for the National Accounts.
- 8.80** A significant improvement would be to separate out current expenditure which invests in programmes expected to have an impact lasting longer than that year, such as modernising the teaching profession or Excellence in Cities. This would improve analysis of the productivity of input expenditure, on the proposed triangulation approach.
- 8.81** Publication of a productivity article along the lines proposed in paragraph 5.40 would provide an opportunity to examine input, output and productivity in greater detail and in context. This could include changes in staffing and resources, and improvements in quality and achievement measured through other means such as those which underpin the PSA targets. Pupil attainment could be placed in context and compared with inspection reports, school resources and teacher quality. Expenditure in schools could be linked more accurately to achievements where there is a lagged delivery. Pupil key stage results need to be linked to inputs over several preceding years. It would also be an opportunity to analyse which input expenditure is intended to buy capacity for the future rather than immediate outputs.

### Summary of Work in Progress and Future Directions

- 8.82** During the rest of the review, we will work with ONS, DfES and the education departments of the devolved administrations to undertake the following work.
- a) We will examine whether a volume measure can be produced that is closer to pupil hours taught than the current measure of full time equivalent pupils. This would look at length of school days and absence rates.
  - b) We will explore further two approaches to the measurement of quality in education as part of value added by schools: the measure of pupil progress based on attainment through examination results; and on a measure of the quality of teaching, or overall school performance, derived from Ofsted inspections. If further work suggests both are viable, they could be weighted together in a value-added measure: if so, we would need to consider carefully what their relative weights should be.

- c) We will seek additional measures of school examination attainment post-16, since this has been an area of increased expenditure not matched by current proposed measures of attainment.
- d) We will consider whether evidence of the value of education which reflects the wider economy, such as net present value of future earnings and the benefits of child care provided by schools, could be used in the National Accounts.
- e) We will examine further the new output measure introduced for health professional education and consider what improvements might be made, including ways of incorporating quality change.
- f) We will consider whether there are other areas where government expenditure on education is not at present classified as such and, if so, what specific output measures should be used.
- g) We will improve the analysis of input data by identifying current inputs that are intended to build capacity for the future or can only deliver outputs after a lag, and improvement to pay and price indices.

**8.83** We are conscious that our work is not covering the higher or further education sectors at all, since these are by definition outside the government sector of the National Accounts and so outside our terms of reference. In the longer term, given the importance of education to the economy, we feel it would be worthwhile for ONS to look at whether its current measures are capturing the added value of this sector, perhaps as part of the suggested satellite account on formation of human capital.

# 9 Public Order and Safety: A Progress Report

---

## Introduction

**9.1** Under the international classification of the functions of government used for compiling the expenditure measure of GDP, Public Order and Safety has six subsections. These are:

- Police;
- Fire;
- Law Courts;
- Prisons;
- Research and Development in Public Order and Safety; and
- Public Order and Safety not elsewhere classified (Probation for our purposes).

**9.2** Three UK government departments take principal responsibility for the Public Order and Safety function:

- Police, Probation, Prisons and Research and Development are the responsibility of the Home Office.
- Fire is the responsibility of the Office of the Deputy Prime Minister (ODPM).
- Courts is the responsibility of the Department for Constitutional Affairs (DCA).

Other departments, such as the Crown Prosecution Service (CPS), also play an important role.

**9.3** In all of these areas, there are distinct legal and structural arrangements for Scotland and Northern Ireland. Services for Wales follow the same systems as in England although most of them are the responsibility of the Welsh Assembly. The exception is that police forces in Wales are the responsibility of the Home Office.

**9.4** It is not straightforward to separate out the various functions of Police, Law Courts and Prisons and tidily quantify their contribution to economic output. The effectiveness of each agent of the criminal justice system (CJS) depends, to varying degrees, on the effectiveness of the others. This is reflected in the government setting some Public Service Agreements (PSAs) for the CJS as a whole, which the Home Secretary shares with the Lord Chancellor and the Attorney General. Each department has its own exclusive targets, but joint ones from 2002 include:

- reduce crime and the fear of crime;
  - improve the delivery of justice by increasing the number of crimes for which an offender is brought to justice;
  - improve the level of public confidence in the criminal justice system; and
  - protect the public and reduce reoffending by five per cent.
- 9.5** In addition to these departmental aims, objectives and targets, a variety of policy initiatives and management frameworks are in operation or being developed and can be used in assessing the services provided by the various departments.
- 9.6** At the Home Office, the Policing Performance Assessment Framework (PPAF) has become fully operational this year (2004-05). There is another, less well developed, initiative to take the processing of offenders away from the prison service and the probation service as separate departments, and put it under a new National Offenders Management System, which will have overall responsibility. A Comprehensive Performance Assessment for individual fire and rescue authorities (FRAs) is being developed.
- 9.7** Although these various performance measurement frameworks (including PSAs) are not directly transferable to a National Accounting framework, they provide an important theoretical anchor for thinking about performance as well as being a driving force behind a great deal of the relevant data collection and analysis. So it makes sense from both a conceptual and a practical point of view to consider using these as the building blocks for the relevant output measures in the National Accounts.
- 9.8** As discussed in paragraph 2.18, it is not easy to ensure measures of public sector output are invariant to changes in the organisation of delivery. Policy initiatives may alter the expenditure priorities of the government and the nature and quality of the services it delivers. Problems are reduced if output measures rely, at least in part, on outcomes, since the ends of government may well enjoy greater constancy than do the means to achieve them. And, importantly for the purposes of this review, they are also less variable over time.
- 9.9** To offer just one striking, and relevant, illustration of this, the outcome measures of police activity in PPAF are:
- reducing crime;
  - investigating crime;
  - promoting public safety; and
  - providing assistance.

**9.10** In 1829, Sir Richard Mayne, one of the Commissioners charged with setting up the Metropolitan Police Force, wrote:

‘The primary object of an efficient police is the prevention of crime [target 1]; the next that of detection and punishment of offenders if crime is committed [target 2]. To these ends all the efforts of police must be directed. The protection of life and property, the preservation of public tranquillity [targets 3 & 4], and the absence of crime, will alone prove whether those efforts have been successful and whether the objects for which the police were appointed have been attained.’

**9.11** So the objectives of the police and its perceived role in society are very similar now to what they were 175 years ago. But the way the police went about achieving these ends in 1829 and the resources they would have employed would have been so different from those of the present day as to make any close comparison very difficult indeed.

### How UK Public Order and Safety Output is Currently Measured

**9.12** In the UK National Accounts, all Public Order and Safety, except for Police, is estimated using activity indicators, without quality adjustment. Police output is measured using deflated inputs.

#### Police

**9.13** Although currently measured using deflated expenditure, some experimental output indicators had been constructed for Police by ONS prior to the initiation of this review. They divided the work into:

- crime related incidents;
- non-crime road incidents; and
- other non-incident related activities, such as patrols, crime prevention and special operations and events.

**9.14** The output indicators for crime related incidents and non-crime road incidents were fairly comprehensive in their coverage but, for lack of data, indicators for the third area did not cover all activities. There was also no proposed method for quality adjustment in any of the three areas.

**9.15** The proposed weights to be used for each indicator reflected the relative cost of clearing up a crime, taken from a survey conducted of the Humberside Police force, although with an adjustment for the Metropolitan Police. This was necessitated by data limitations, but the assumption that one police force was representative enough to use for the construction of output weights would seem questionable.

### Courts and Administration of Justice

- 9.16** These are currently estimated using direct output methods. The output of the county courts, the crown courts and the magistrates courts are measured separately, with the workload broken down by type of case. For county courts, there is also a measure of the administrative workload, measured by number of hours. The weights for each type of category are calculated on some measure of average time/average cost, depending on data availability.
- 9.17** Outputs for the Crown Prosecution Service and the Legal Services Commission (legal aid), and their corresponding weights are calculated in similar fashion, with a workload measure of output and an average time/cost weighting system.

### Probation

- 9.18** This is currently estimated using direct output methods. The current indicators used for Probation are probation starts; community service; combination orders and licences; numbers of pre-sentence reports (PSRs) completed; and a measure of probation work done in the family court. The Home Office's cost-weighted activity index is used to construct output weights for these various services.

### Prisons

- 9.19** These outputs are currently measured by numbers of nights in prison of prisoners on remand; prisoners under sentence; non-criminal prisoners; and prisoners in police cells. They are unweighted.

### Fire

- 9.20** Fire output is currently measured directly. There are three broad output categories. These are:
- fighting fires;
  - preventing fires; and
  - special services.
- 9.21** The categories are further divided into sub-categories, or incident types, and output weights for these are assumed to be proportional to average staff hours spent on each one. The total weight of each incident type is calculated as the product of the weight for one incident and the number of such incidents in the base year. For prevention, the output measure is staff hours.

## International Context

- 9.22** The Eurostat Handbook rules on A/B/C methods in the non-market sector for collective services are not as stringent as for individual services (see Appendix C, Table C3).
- 9.23** The ESA provides certain guidance/ground rules for the measurement of collective services. For non-market services in general (i.e. individual as well as collective) it stipulates that the production and consumption of goods and services must be defined by their flows and not the final results obtained by their use. But the guidance also says specifically of collective services that the volume of their output cannot be captured by the extent to which they are utilised (see Appendix C). In other words, this suggests that it is not desirable to measure the output of Fire by the number of calls attended (as the United Kingdom does now, see paragraphs 9.20-9.21), since it is in the availability of the service that the value lies. Similar parallels could be drawn with Courts and Police.
- 9.24** But the ESA also implies that outcome measures cannot be used because they will be influenced by extraneous factors; and that activity based measures – say the number of fire calls attended, or the number of arrests made – should not be used either, since they do not account for the indirect and continuous nature of a collective service.
- 9.25** To square this circle and keep within the spirit of the guidelines, the challenge must be to attempt to find indicators that, if they are outcome measures, can be adjusted to strip out the effects of any extraneous factors; or, if they are activity measures, can be adjusted to give an indication of added value to society of the service concerned.
- 9.26** No European country uses a direct volume measure for Public Order and Safety output in their national accounts.
- 9.27** The Australian Bureau of Statistics (ABS) has done some work in recent years on developing output measures for police services, justice services and corrective services. Its initial findings were published in a report in November 2000. An output measure for police was constructed, drawing primarily on ABS data on recorded crime. However, the measure was not considered suitable for implementation due to the number of assumptions required for the index to be representative of the sector. The report recommended that an input-based measure be retained for Police. It also suggested the kind of data and methodology that would be required in future for a credible output-based measure of police services. Since then, a decision has been made to keep with the deflated input methods. Further details are in Appendix D, paragraphs D51-53.

### Critique of Current Public Order and Safety Output Methods

- 9.28** The Public Order and Safety services are, for the purposes of National Accounts, considered to be collective services. The *System of National Accounts* (SNA) summarises collective services as those that can be delivered simultaneously, do not require the explicit agreement or active participation of all the individuals concerned and for which there is no rivalry in acquisition (see Appendix C, paragraph C14).
- 9.29** Clearly, some of the services combine elements of both collective and individual services. Arguably, there may be rivalry in acquisition for the fire service, the police service, or the use of the courts. A fire engine, when on active service, must be devoted to servicing individuals or individual buildings. There may be times when the demand for fire engines in a locality exceeds their supply. But the fire service is a collective service in so far as the intention of the authorities is to provide sufficient capacity to meet all demand. The purpose of the service is that it should be universally available, within a reasonable response time, wherever and whenever the demand arises, in the same way as street lighting.
- 9.30** A similar point could be made about the police service. The aim of the service is to provide the public with universal and continuous protection. But, in the event of a failure of this protection, the police also provide a service in clearing up crime, and this is delivered to the individual victims. Here, there will often be rivalry in acquisition. So, for instance, police attendance at the scene of a domestic burglary is often far from immediate. Victim support services are clearly individual, with it possible in principle to measure quality of service received by the person concerned.
- 9.31** The CJS as a whole is best seen as a collective service. The police, prisons and other areas of the system involve the arrest, prosecution and incarceration of individual offenders, but the service being provided is a collective one to society as a whole, rather than an uninvited one provided to those prosecuted, arrested or incarcerated.
- 9.32** The civil responsibilities of the courts are different, since the services they provide are provided to the individuals or companies involved. Clearly a divorce settlement, or a compensation claim, or a libel case do not meet the SNA criteria for collective services outlined in Appendix C, paragraph C14.

- 9.33** One systemic problem with the current methodologies for the areas of Public Order and Safety that contribute to the CJS is that they take no account of their inter-relatedness. The current system treats the police, the courts, and the prisons as individual, stand-alone entities. But the effective functioning of the CJS requires the processing of offenders from arrest to prosecution, to the delivery of justice – whether punishment or acquittal. An accurate measure of the increment to collective welfare from the criminal justice system should reflect this.
- 9.34** In all areas of Public Order and Safety, including the fire service and those policing services not related to the CJS, there are problems with activity indicators being used that do not correlate well with the value of the service provided. Moreover, in each sector there are areas where clear improvements can be made along the lines set out in paragraph 6.3.
- a) **Capturing all output:** there are areas of output that are not covered by the existing indicators. There are various problems with the quality of the data, particularly those used for the construction of output weights, which are often crude or extrapolated from unrepresentative samples.
  - b) **Improving measures of output to capture changes in quality:** current measures need improvement, with a focus on ways of understanding and valuing the additional collective outcome attributable to expenditure in these areas.
  - c) **Improving UK coverage and timeliness:** most of the output indicators used in Public Order and Safety cover only England and Wales. More work is needed, with the devolved administrations, to consider how to measure output gained from expenditure in Scotland and Northern Ireland.

### Police

- 9.36** As with many other areas of the CJS, the service provided by the police is often most effective when the police's intervention in the lives of the community is minimal. For instance, a police presence that prevented a fight breaking out in the first place, might in many situations be deemed preferable to a response that brought the fight to a conclusion after it had started and that led to arrests.
- 9.37** So, indicators of police performance that are activity-based are not necessarily good indicators of the improvement or otherwise of the service – an increased number of arrests does not necessarily mean the police are doing their job better. Many of the indicators currently used as experimental indicators to measure Police output are open to this criticism.
- 9.38** The police and the Home Office are fully aware of the shortcomings of these activity-based measures of output and generally do not use them. They tend to concentrate on targeting outcomes, such as reducing crime.

**9.39** The output weights in the experimental indicator are grossed up from the activity based costing model of one force – Humberside. But the activity based costing model, from which the Humberside data are drawn, has now been extended to all areas of the country and so it should be possible to improve the expenditure weights to be used in future.

### Prisons

**9.40** Current practice measures the output of Prisons with unweighted prisoner nights. This is a very crude measure for three distinct reasons:

- it fails to quality adjust for overcrowding, reoffending, achievements during prison such as education or drug rehabilitation;
- it fails to weight according to expenditure – high risk/ low risk prisoners;
- numbers of prisoners and length of sentences may be a perverse indicator which would fall if the CJS as a whole performed better.

### Courts and Administration of Justice

**9.41** The current methods for measuring the output of the courts have a number of defects. Some perverse indicators are used. For instance, in county courts administrative hours are used as a volume indicator of output, although they are clearly an input. In recent years, administrative work has become much more efficient, which ought to result in a measured productivity increase. Instead, the output is deemed to have declined. Other areas for review are:

- the lack of comprehensive coverage – for example there is currently no measure of divorces;
- the lack of a distinction between trials and pleas – essential because pleas are far cheaper; and
- quality adjustment, eg measures of victim satisfaction with or support from the courts.

**9.42** For Probation there are no current data. The figures are extrapolated from a short time series, which is now a few years old.

### Fire

**9.43** The fire service volume measures suffer from some of the same shortcomings as those for the police. There are some perverse indicators, where the volume of incidents attended is used as a proxy for increased output – in line with the discussion in paragraphs 5.48-5.56, a good fire service would, arguably, attend fewer fires, because it would prevent them. Certainly prevention has been the central preoccupation of policy in recent years. So, as with much of police activity, simply counting incidents attended (activities) does not provide a measure of value. Related to that is the absence in the current methodology of any adjustment to the added value of the service predicated on the value of the property and lives it protects.

## Way Forward on Output Methods

- 9.44** We are aiming to improve the measurement of government output for the National Accounts by measuring the incremental contribution to individual or collective welfare, as set out in Principle A (see paragraph 5.7). There are two possible impediments to meeting this objective. The first is the lack of data, which could be overcome, given enough time and adequate resources. The second is that there are some areas of government output where it is practically impossible to measure the incremental contribution to welfare of a service, even where it is clearly obvious that the value exists.
- 9.45** For example, it is possible to construct an adequate measure of the police contribution to clearing up burglaries. But around half of police resources are devoted to promoting public safety and other non-crime work. There is clearly a value in this work, but it is difficult to think how to measure that value independently of the inputs.
- 9.46** This causes problems with implementing international guidelines. If, say, 75 per cent of the output of the CJS is covered by indicators that provide a good estimate of its added value, but there is no available estimate of the remaining 25 per cent, we will not be meeting the requirement that output indicators *'should cover all the services produced by the producer that are provided to external users...'* (see Appendix C, paragraph C29).
- 9.47** A decision is needed on what assumption to make about the productivity of the remaining 25 per cent. An assumption of zero productivity (inputs = outputs) appears neutral, though nevertheless arbitrary, when applied across the board. It has no particular merit if applied selectively in an area where there is positive or negative productivity change for other services, since it would shift the overall average productivity with no justification. An alternative, which has effectively been used so far by ONS in areas like Health, would be to extend the productivity estimate of the measured 75 per cent to the unmeasured 25 per cent. Use of such assumptions should be explicit, and, where possible, subject to sensitivity analysis. The problem is greatest where the proportion of services for which added value can be measured is low, or the measure used of doubtful quality.
- 9.48** As the review continues, we will be working closely with the Home Office, DCA, ODPM and other relevant departments and agencies to explore how much of government expenditure on Public Order and Safety can be tied to measurable output, including ways of building quality into the calculation. In addition, we will also be working on whether we can arrive at any independent measures of productivity or efficiency for the remaining areas of expenditure.
- 9.49** In accordance with the earlier discussion, some of the methods described below seek to measure the total impact of the total system, rather than to identify a volume of separate activities and estimate the value of each, as would be the case with a health activity index using quality measures for individual treatments. They could be implemented as a total productivity adjustment to the input value, rather than by summing individual outputs. The implications of this different approach will need careful assessment.

### Output approach to criminal justice system

- 9.50** One possible approach is to use an existing Home Office model that attempts to measure the contribution of the Criminal Justice System to crime reduction – the principal departmental outcome target. The model controls for other non-CJS factors such as the economy, and social and demographic factors. It can be used to show the marginal effect on the level of crime of increases or decreases in CJS activity.
- 9.51** Changes in levels of crime can be converted into money values using Home Office estimates of the economic and social costs of crime. The value of crime reduction can then be compared with the value of CJS inputs – so it allows for a type of cost-benefit analysis. The model is well established, but is subject to continuous development and updating. The Home Office expects to publish the findings of this work in due course.

### Output from quality adjusted activities

- 9.52** This approach would be based on the principle that individual crime reduction delivery agents produce activities that can be quantified in aggregate (using a cost-weighted activity index approach, for example). So the approach would be to quality adjust total activity values by some factor based on total crime reduction, controlling for other external factors that may be affecting this outcome.

### Output modelling based on offenders brought to justice

- 9.53** This approach tries to define the output of the CJS as a whole. In approach, it is similar to that in paragraph 9.52, but focuses on offenders brought to justice and acquittals, rather than a broader set of outputs, attributable to individual agents. The overall output with this model is ‘the effective resolution of an incident’, resulting in an offender being acquitted, or sanctioned in some way. It should be noted that sanctioning is not solely down to the courts; the police, through cautions and penalty notices, also sanction offenders.
- 9.54** This method would treat crime reduction, public disorder, traffic safety, the effective conclusion of civil court cases and other non-CJS related outputs separately.

### Comment

- 9.55** Each of these methods has its merits. Whichever is chosen as the central approach, the information within, and approach, of the other two methods could be used to improve, or cross reference results. For instance, the information provided by the model outlined in paragraph 9.52 could inform the required productivity assumptions that will be needed to quality adjust the approaches outlined in paragraph 9.50-9.51, or to cover areas where output indicators are not clearly identifiable.

### Civil courts

- 9.56** As noted in paragraph 9.32 the work of the civil courts is closer to an individual service than a collective one. Finding indicators of output should as a result be somewhat easier. Work to improve current methods will concentrate on ensuring that all output is captured, dropping the use of input measures as a proxy for output, ensuring that expenditure data for the construction of weights is improved, and that a robust measure of quality change can be developed.

### Fire

- 9.57** The Fire and Rescue Service Improvement Team (FRSIT) is working on a suite of indicators to capture efficiency/ productivity in the fire service that captures all these various methods of prevention mentioned above. This work is independent of our review, but it is likely to inform any practical way forward.
- 9.58** It should also be possible, in the medium term, to make use of the fire service's Comprehensive Performance Assessment (CPA). This will start in January 2005 with a view to having a report on every English Fire and Rescue Authority (FRA) by August 2005.
- 9.59** In addition, ODPM have developed a 'cost of fire' model. This seeks to define and calculate the cost of fire to England and Wales, including how much it costs businesses to adhere to safety regulations and how much households spend on smoke alarms. It also tries to quantify how much is spent on community fire safety, the administrative costs of insuring properties and the costs of lost property, lost business and loss of life as a result of fire.
- 9.60** These initiatives, together with the 'Integrated Risk Management Plans' of each FRA will be used by the review team to explore the possibility of constructing a robust measure of Fire productivity.

## Inputs and Deflators

### Inputs

- 9.61** Expenditure estimates from a number of different central government departments, by local government and by the devolved administrations, are combined into the COFOG category of Public Order and Safety.
- 9.62** Total current and capital spending for 2004/05 was £17bn for current spending and £1bn for capital. Expenditure on the Crown Prosecution Service and DCA would add around another £1.5bn to this figure, bringing the total to about £20bn. To measure productivity, it will be important to examine the breakdown of these figures in detail. Current expenditure will include items such as staff training that, while not capital items in a strict accounting sense, are creating capacity for the future and should not be expected to deliver immediate outputs.

### Deflators

- 9.63** Improving the collection of current price data on expenditure will only be of any benefit if the deflators used to produce volume measures are of equal precision, detail and coverage. Most of the areas covered by Public Order and Safety use an average earnings index for the government sector to deflate pay in central government expenditure, and a similar index for the local government element. The exceptions are that local government compensation of employees for Police and Fire are deflated by specific local authority pay indices. Local government expenditure accounts for around 40 per cent of all police expenditure and almost all fire service expenditure. It would be desirable to develop specific deflators for these areas.
- 9.64** The methods used for deflating procurement use a combination of Producer Price Indices and Retail Price Indices. This should be adequate, but ONS does not receive expenditure data in sufficient detail to be able to deflate this expenditure accurately, or attach the correct weights to the various deflators. Similar difficulties exist with the deflators for capital consumption. Attention needs to be given as to how best to tackle these problems.

### Triangulation and Productivity Measurement

- 9.65** The implication of the international guidelines discussed above is that the output of collective services should be assessed on their quality and their availability and that these are more likely to be linked to outcomes, such as a fall in crime, than to the extent to which the service is utilised. Or, looking at the problem in terms of the triangulation process in Principle G (see paragraph 5.35) for these services, independent evidence on productivity performance is likely to be at a particular premium.
- 9.66** The ESA seems to suggest as much when it says:

‘...For collective services it is generally not possible to establish unit costs and quantities reflecting their utilisation. If attempts are made to account for changes in productivity for collective services by indirect methods, users should be made aware of this.’ (ESA, 1995, paragraph 10.25.)

### Summary of Work in Progress and Future Directions

- 9.67** Our main concern is to ensure that the measures of a collective service reflect the added value to society, rather than simply measuring volumes of activities. This may best be achieved by looking at overall outcomes rather than attempting to ascribe quality to individual activities. For the civil courts, however, an approach to valuing individual activities for individual end users is more applicable.
- 9.68** During the rest of the review we will:
- a) work closely with the Home Office, DCA, ODPM and other relevant departments and agencies to explore how much of government expenditure on Public Order and Safety can be tied to measurable output, including methods of building quality into the measurement;

- b) consider a triangulation approach using independent productivity measures or measures of the effectiveness of the whole system, where expenditure does not match activities whose value added can be measured;
- c) work with the Home Office and other departments on the alternative approaches to measuring output discussed above;
- d) work with DCA to improve measures of the civil courts' output, ensuring all output is captured, that expenditure data for the construction of weights is improved and a robust measure of quality change is developed;
- e) explore whether a number of different measures of fire service productivity could be used to assess the overall value added by the service, with a triangulation approach to be considered as a possible improvement on current activity measures;
- f) work with ONS and the relevant departments to ensure expenditure data are complete for each function, and are correctly assigned to relevant outputs;
- g) identify aspects of expenditure, like staff training, that should not be expected to lead immediately to extra activity and consider how to take account of this expenditure in developing and interpreting productivity measures;
- h) work with ONS and departments to ensure satisfactory deflators are available for each expenditure area; and
- i) work with devolved administrations towards using direct data, having regard to the separate systems in place in the various countries, to underpin the measure of UK output.

# 10 Social Protection: A Progress Report

---

## Introduction

- 10.1** Social Protection consists of Personal Social Services and the Administration of Social Security.

## Personal Social Services

- 10.2** Personal social services comprise mainly the provision of benefits in kind in order to improve, or prevent the deterioration of, the lives of certain individuals and their carers. These are provided where individuals are vulnerable because of age, infirmity, mental health problems, learning or other disabilities; or because they are families and children in vulnerable circumstances. Examples of personal social services activities include: residential care provided to vulnerable adults and children; day centres for older people and younger adults with learning disabilities; home care and help services; meals on wheels; providing equipment and adapting homes to facilitate independent living; placements into adoption or fostering of ‘looked after children’; assessment of the needs and the care management of vulnerable individuals; maintenance of child protection registers and the associated casework.
- 10.3** Local government is responsible for the delivery of personal social services. Councils are responsible for assessing the needs of their populations and for arranging care in the light of local priorities and national guidance. They purchase care from both public sector providers and voluntary and private providers. The greater part of formal care services for adults is provided by independent sector providers – that is, voluntary organisations or private companies. These providers generally supply services both to publicly funded clients and to clients purchasing care privately.
- 10.4** Eligibility for publicly funded social services for adults depends on an assessment of care needs and, for most care services, a financial assessment. The latter takes account of the individual’s income and savings. Individuals with assets above a capital limit are not eligible for publicly funded care. They may purchase services privately and, in some cases, may move from privately funded to publicly funded care as their savings are depleted.
- 10.5** Provision of social services is a devolved responsibility in Scotland, Wales and Northern Ireland. The financial arrangement in Scotland is different from the rest of the United Kingdom. Personal care for older people in Scotland is not subject to means test, but other forms of care are.

**10.6** Local government provision of social services is funded by grants from central government (including the devolved administrations) and by local taxation (council tax). In England, the Department of Health (DH) is responsible for social services for adults, and since 2003, the Department for Education and Skills (DfES) has been responsible for children's social services. DH is responsible for matters relating to the social care workforce. The Office of the Deputy Prime Minister (ODPM) is responsible for the funding of local authorities and the framework of financial and performance accountability of councils in England. Other government agencies, including the Audit Commission and the Commission for Social Care Inspection, also play important roles.

**10.7** The 2002 Public Service Agreement set the following objectives for social care:

- a) improve life outcomes for adults and children with mental health problems, through increased access to services;
- b) improve the quality of life of older people so that they can live at home wherever possible;
- c) narrow the gap in educational attainment between children in care and their peers;
- d) narrow the gap in criminal convictions and cautions between children in care and their peers; and
- e) improve value for money in NHS and personal social services. This is to be assessed on:

'unit costs of procedures and services, adjusted for quality, underlying inflation and mix of cases. The service effectiveness element of the target [is to be] based on quality indicators published by the department'.

**10.8** A fuller statement of social services objectives is provided in 'Social Services Performance Assessment Framework Indicators 2002-03', published by DH. Many of the headings have an associated 'Performance Assessment Framework' (PAF) indicator, under the PAF system introduced in 2000 to improve performance assessment by local authorities with social services responsibilities. The indicators take a variety of forms such as: unit cost measures of inputs; quantitative measures of achievement of outcomes; and subjective client satisfaction surveys.

**10.9** There are a number of conceptual difficulties in identifying and measuring the incremental value added by social services. In this area, added value might be considered by comparison with the counterfactual of what would happen with no service – often, deterioration, disability or further loss of dignity and autonomy. The purpose of these social services is to maintain or improve the quality of lives of individuals who are vulnerable or in need in some way or other. An increment to welfare will depend both on the starting point, eg particular severity of disability, and the outcome, eg a capacity to live independently at home. Not all the increment will be attributable to the intervention of social services: some will depend on external circumstances or on the impact of informal care (eg by family members), or may be due to other public services such as health care.

- 10.10** Social services for adults help people who experience difficulties with personal care or other key tasks of daily living. Their aim is to improve quality of life across a number of dimensions that could include independence, dignity, choice, or social interaction. Such improvements are not straightforward to measure. For some, the prevention of further loss of independence is itself an outcome.
- 10.11** Many recipients of social services also receive informal care or other services, such as health or supported housing, that also contribute to improvements in quality of life. This complicates the attribution of improvements in quality of life to social services.
- 10.12** It is important to recognise that the greater part of care for adults is provided by unpaid informal carers. Respite and other support for informal carers is a legitimate output of social services, in part because sustainable informal care avoids the need for more expensive, paid care, and also because it meets the needs of informal carers, who may themselves be in poor health. (Informal care, like other unpaid arrangements within households, by convention does not feature in the National Accounts, whatever its value to the individuals concerned.)

### Social Security

- 10.13** Compulsory social security within Great Britain is a reserved responsibility of the Department for Work and Pensions (DWP). There are parallel arrangements in Northern Ireland. The social security system delivers social welfare transfers to eligible categories of recipients. These include: the state basic retirement pension and related schemes; job-seekers' allowance and income support payments payable to economically inactive people; child support maintenance payments where parents are separated (including both government payments and mandatory private transfers); housing cost subsidies (housing benefit); and the operation of certain hybrid tax and benefit schemes designed to preserve positive incentives for workforce participation and private saving.
- 10.14** Social security functions in the United Kingdom also include the provision of benefits in kind to eligible individuals. These include employment advice and training provided to unemployed people; programmes aimed at reducing child poverty and pensioner poverty; and programmes designed to promote the participation of disabled people in work and training.
- 10.15** DWP's activities are provided through its executive agencies and, in the case of the administration of housing benefit, by local authorities with most costs reimbursed by DWP. The social security functions of paying child benefit and collecting compulsory social security contributions are now the responsibility of the Inland Revenue, as is the administration of tax credits which have replaced some former social security benefits.

**10.16** Within social security, there has been a considerable programme of change in organisational structures, new forms of benefits and tax credits, investment in staff and premises, and new objectives for delivery of support to individuals. DWP itself was formed in 2001 during machinery of government changes, and has been internally restructured. Programme changes include reforms of the structure of tax credits and benefits, replacement of state earning related pension scheme (SERPS) by the second state pension, and continuing changes to the nature of the services offered by Jobcentres to unemployed people in various categories. These are associated with a three-year programme of enhancement of Jobcentre facilities.

**10.17** The DWP's key delivery objectives are listed in its 2004 departmental report. These are:

- a) ensure the best start for all children and end child poverty in 20 years;
- b) promote work as the best form of welfare for people of working age, while protecting the position of those in greatest need;
- c) combat poverty and promote security and independence in retirement for today's and tomorrow's pensioners;
- d) improve rights and opportunities for disabled people in a fair and inclusive society; and
- e) modernise welfare delivery so as to improve the accessibility, accuracy and value for money of services to customers, including employers.

**10.18** Associated with these objectives are ten specific Public Service Agreement targets, some of which are shared with other departments, relating to:

- a) quantified reductions in child poverty over specified horizons;
- b) increasing the proportion of parents with care on income-based benefits who receive maintenance;
- c) increasing the employment rate and reducing the unemployment rate over the economic cycle;
- d) increasing employment rates for specified disadvantaged areas and groups;
- e) reducing the proportion of children in households without work;
- f) payment of Pension Credit to pensioner households;
- g) increasing the employment rate of disabled people, conditional on the economic cycle, and improve the rights of the disabled;
- h) proportion of benefit payments paid directly into bank accounts;
- i) publication of demanding departmental performance targets; and
- j) reduction in fraud and error losses.

**10.19** If social security is narrowly defined as the provision of monetary transfers, then a measure of value added could be defined as a volume index of numbers of benefit claims processed, which could be adjusted for quality by reference to indicators of accuracy, timeliness, control on fraud and other relevant criteria. But this premise has become increasingly unrealistic, as DWP's objectives have developed to include the active promotion of workforce participation, security of pensioners (including through private pension schemes), and similar themes. A broader scope of output measurement is needed to capture the full range of current social security activities.

### How UK Social Protection Output is Currently Measured

**10.20** Direct output measurement of UK personal social services and social security was introduced in its current form in 1998, backdated to 1994. There is roughly a year time lag on the publication of data.

#### Personal Social Services

**10.21** Personal Social Services output is measured in the National Accounts by a cost weighted index of selected activity indicators, with weights based on available net expenditure data compiled from local authorities' statistical data returns to central government. Expenditure data are collected on both a gross and a net basis, where the latter basis deducts charges made to individuals.

**10.22** Coverage of actual activities is far from complete, and cost weightings used in ONS's compilation are multiplied up so that they sum to 100 per cent of expenditure on the net measure (see Table 10.1). Table 10.2 summarises the composition of the cost weighted activity index for Personal Social Services.

**10.23** The measured activities currently in the index are:

- numbers of individuals accommodated by the local authority, broken down by children and adults, and by available categories of residential placement; and
- the volume of home help – home care support provided daily to older adults living at home, measured in hours.

**10.24** Total residential care accounts for approximately 77 per cent of the total weighting, within which the greater part is for older people. Home help accounts for 23 per cent. These weightings are roughly double the respective shares that these activities represent in net expenditure.

Table 10.1 **Personal Social Services: England local authorities net expenditure. Weights in ONS output index 2000/01**

Heading / subheading	£ billion	Share of total net expenditure (%)	Output indicator weighting (%)
A: Service Strategy	0.1	1.0	
BA: Children's services – commissioning and social work	0.8	7.4	
BB: Children looked after	1.3	12.2	28
BC: Family support	0.4	3.8	
BD: Youth justice	0.1	0.9	0.2
BE: Other children's and family services	0.2	2.2	1.4
C: Older People	4.3	39.9	70.4
<i>Of which:</i>			
<i>residential</i>	2.2	20.6	47.1
<i>home care</i>	1.1	10.2	23.3
<i>other</i>	1.0	9.1	
D: Adults under 65 with physical disability or sensory impairment	0.8	7.1	
E: Adults under 65 with learning disabilities	1.5	13.7	
F: Adults under 65 with mental health needs	0.6	5.6	
G: Asylum seekers	0.5	5.1	
H: Other adult services	0.1	1.0	
<b>Total Personal Social Services</b>	<b>10.7</b>	<b>100.0</b>	<b>100.0</b>

Source: DH and ONS

Table 10.2 **Activity components in ONS Personal Social Services index: levels and weights 2000/01**

Sub-Index	Number (000s)	Weighting (%)
Adults in residential care, <i>of which:</i>		47.1
<i>Local authorities</i>	44.8	20.6
<i>Independent residential care home</i>	140.3	13.8
<i>Independent nursing homes</i>	72.9	12.8
<i>Unstaffed, etc.</i>	5.7	0.0
Home helps (hours)	2,836	23.3
Children in residential care at 31 March, <i>of which:</i>		29.6
<i>Foster placements</i>	38.1	11.7
<i>Lodgings</i>	1.2	1.4
<i>Community homes</i>	4.8	6.7
<i>Voluntary homes</i>	0.7	2.8
<i>Private registered</i>	1.2	1.3
<i>Schools, etc.</i>	1.1	1.4
<i>Placed for adoption</i>	3.3	0.8
<i>Placed with parents</i>	6.7	2.8
<i>Other accommodation</i>	1.8	0.6

Source: DH and ONS

### Social Security

**10.25** The National Accounts classify social security payment flows as transfers of income and the administration of such payments as government output. The output of social security services is measured as the cost-weighted index of volumes of new benefit claims processed, broken down by principal benefit programme. Cost weightings derive from unit expenditures in each programme. Weightings are set out in Table 10.3.

Table 10.3 **Components of Social Security output index**

Indicator (number of new claims)	Weight %
Retirement Pension	3.1
Widows Benefit	0.5
Jobseeker's Allowance	2
Sickness Benefit	5.1
Income Support	58.2
Family Credit [ceased]	–
Social Fund	7.9
Other Benefits	5.8
Housing Benefit	17.3
<b>Total</b>	<b>100</b>

Source: ONS

### International Context

**10.26** Social Protection services are classed in the main as 'individual services', since they operate to the advantage of identifiable individuals. As such, the European Commission decision requires measurement to be based on output indicators by 2006. Output indicator approaches that are detailed, comprehensive, encompass quality change and are accurately weighted are A, or ideal methods. Output indicator methods without quality adjustment may be B if otherwise satisfactory, or C if deficient in other respects, like lack of detail. Inputs-based methods are classed as C, or unacceptable methods. Some components of Social Protection services, such as the promotion of opportunities for disabled people, are classified as collective services for which the use of inputs measures would be a B method.

### Personal Social Services

**10.27** Guidance in the Eurostat Handbook covering social care services distinguishes between those including and not including accommodation, and recognises that delivery of such services may be in a market or non-market context. The social care element of social care services should be separately identified from the nursing element, which should be classified to Health, although it is accepted that, in practice, this may be difficult to achieve.

**10.28** The Handbook advises that, for the market provision of social care services, an A method is to deflate expenditure by an appropriate price index. For non-market provision of social care involving accommodation, measurement by occupant days broken down by type of institution and fully adjusted for changes in quality will also qualify as an A method. Occupant days by themselves will qualify as a B method. For non-market provision of other social care services, that is, not including accommodation, the recommended approach for an A method is measurement by numbers of persons receiving care, broken down by levels of care. Total numbers receiving care will qualify as a B method.

### Social Security

**10.29** The Handbook offers little guidance on the measurement of social security, except to suggest that numbers of benefit recipients, adjusted for quality by reference to factors such as speed of processing of applications, and numbers of errors, might serve as a starting point an output indicator. The Handbook includes social security administration within a wider class of public administration and defence services, where it recognises that progress towards direct output measurement is especially difficult.

**10.30** Few other countries have yet adopted direct measurement approaches for Social Protection. In certain countries with insurance-based regimes for social care, social care with accommodation is measured by deflating expenditures by standard insurance rates for accommodation.

### Critique of Current Social Protection Output Methods

**10.31** The current methods of output measurement for social services and social security are discussed below in terms of the criteria in paragraph 6.3.

### Personal Social Services

**10.32** The current output indicators correspond to 44 per cent of total social services expenditure. They omit any measure of activities to help adults under 65 with mental health problems or learning or physical disabilities, clearly areas where social services do make a difference to individuals and their families. There are no measures of activities to support children other than residential placements; eg casework on child protection and support for care leavers are not covered. There is no coverage of social care activities that have a 'throughput' dimension, such as assessments of need, including casework relating to the discharge of patients from hospital who cannot cope without support, or the assessment and fitting of home adaptations for people with disabilities. Overall, it does not look safe to assume that changes in outputs for the measured 44 per cent of activities are an adequate proxy for all other areas of adult social services.

- 10.33** Some 77 per cent of the volume measure is represented by residential care of children or adults. Table 10.2 shows the detailed coverage of different types of residential accommodation, by accommodation provider for children and adults. Twenty-one per cent of the total social service indicator is for residential care for adults provided by local authorities. There is no breakdown by level of dependency or type and quality of accommodation provided, such as whether residents have single rooms. The measure is a census of resident numbers in different placement types on 31 March each year. This will represent total occupant days only if there are no significant variations in resident numbers or between types of placement within any given year.
- 10.34** There is no quality dimension to the current social care measures. This is a major concern, given our desire to measure added value of services which may vary in the extent to which they meet the needs of individuals and improve the quality of their lives.
- 10.35** There is no breakdown by dependency or other needs factors in the current measure. There is evidence that the average dependency of older care home residents has risen in the past and that home care services for older people have become more concentrated on those with higher dependency. Data on numbers of clients do not take account of any changes in intensity due to changes in client needs. More dependent residents of care homes may require more hours of care per week than less dependent residents.
- 10.36** ONS uses data for England as a proxy for UK output. Data are updated once a year, based on measures taken by local authorities on 31 March each year. Improvement in both areas looks to be desirable.

### Social Security

- 10.37** Social Security output is currently measured only in terms of new claims received. This method has a narrow scope. It fails to encompass the wider welfare objectives of DWP, such as getting people into work, reducing child poverty, or increasing pensioner saving.
- 10.38** Different new claims are detailed by type of benefit. But there is no count of work done to continue payment of existing claims, including, for example, measures of queries, changes in beneficiaries' details, appeal decisions, fraud checks or annual updates in payment rates, although data about these activities are available within DWP. It is unlikely that ongoing costs are a constant ratio of the costs of handling new claims for each type of benefit, as rates of churn vary considerably for different benefits. A more detailed count of ongoing work would add to the validity of the overall measure.
- 10.39** The current measure fails to register important dimensions of quality of service such as accuracy of claims, turnaround time, and the reduction of fraud. Nor does it gauge whether DWP is adding value to the wider objectives, such as getting people back to work.

- 10.40** The current measure covers Great Britain only and does not use any data on outputs from the Northern Ireland social security system.
- 10.41** Claims and other data which can be derived from the DWP's administrative systems ought in principle to be available frequently. More investigation is needed on current links between ONS and DWP on timeliness. However, other data, for example on fraud and housing benefit, are less timely and less frequent.

## Way Forward on Output Methods

### Personal Social Services

- 10.42** DH has been working on the measurement of the efficiency and quality of adult social services for some time. It is pursuing the construction of an interim social services output measurement index that will also evaluate the delivery of effective social services as required by the 2002 PSA target for improving value for money. This will be an improvement over the current output series used by ONS because the coverage of activities will be more complete. It would contain measures of activity for younger adults, including those with learning disabilities, physical disabilities and mental health problems. It will also cover a wider range of non-residential services by including day care and meals. Assessments would be counted as an output, since they contribute directly to individuals' welfare by giving them advice and information about their entitlements or care plans. DH expects that the new measure should cover activities relating to some 90 per cent of expenditure.
- 10.43** DH plans to discuss their proposals with interested parties and hope to be able to provide cost-weighted data for the new output series to ONS by autumn 2004. Back-dating will be complicated by discontinuities in data and changes in responsibilities. In particular, local authorities assumed from DWP in April 2002 the responsibility for funding a cohort of some 50,000 care home residents who had had preserved rights to income support payments by virtue of admission before April 1993. Although we feel that this would be a useful step forward, it is not a complete solution because it does not measure quality (or dependency).
- 10.44** DfES is also considering whether improvements can be made in the coverage of the output series on children's social services now used for the National Accounts. This may take account of a new focus on outcomes for children following from the machinery of government changes and the Green Paper, *Every Child Matters*, published in September 2003.
- 10.45** Assessing added value by reference to individuals' welfare requires information on quality of social care provided, preferably from the perspective of the individual receiving it. It requires a basis for attributing outcomes to the social care intervention, stripping out the effects of external factors or individual developments like ageing.

- 10.46** DH is funding academic research on the measurement of output and productivity in adult social care. We hope that the discussion in the report will help focus that research.
- 10.47** Performance data, to be useful as output activity indicators, will need to satisfy criteria of timeliness, accuracy, and relevance to the measurement of quality-adjusted final services received by an individual. The development since 2000 of the Performance Assessment Framework (PAF) and the Referral, Assessment and Packages of Care (RAP) indicator frameworks for social services provide some promising concepts and source data for taking this forward.
- 10.48** Collection of new data on a national basis would add to costs of data recording and analysis. This is an area where research evidence or periodic sample surveys could be justifiably used and more practical than seeking to measure quality for every individual social care episode. But there may be practical and ethical issues involved in sampling measures of quality of care for some social services clients, who might, for instance, be very confused or disturbed.
- 10.49** It is important to extend the measurement of outputs to take proper account of data and analysis from Scotland, Wales and Northern Ireland. We will be working further with the devolved administrations during the remainder of the review.

#### Social Security

- 10.50** We are working with DWP to identify possible changes in the current measure of social security output, based on understanding the full range of DWP's objectives and the use of resources. Some questions are issues of sectorisation. For example, assessing how much is spent, and what is achieved, in a particular welfare programme. But others require the exercise of accounting judgement in distinguishing between spending that supports immediate outputs and spending that builds capacity for the future, such as elements of the change programme.
- 10.51** DWP has identified a range of candidate output measures and has begun to assess the availability and quality of data. For employment, where outputs are not currently measured in the National Accounts, there is a range of robust activity indicators which could be considered, such as job entries, vacancies filled or work-focused interviews conducted. These are included in the Jobcentre Plus Performance and Resource Agreement, and performance data are produced on a regular basis. Construction of output indicators for other areas, such as the DWP's promotion of private pension saving or of activities that reduce fraud and error, may be more difficult.
- 10.52** DWP has also begun to consider how outputs should be weighted, and is investigating the suitability of a recently introduced activity-based costing system. This is likely to be more appropriate than the current weighting system. It will also consider how quality should be taken into account, for example through measures of accuracy or added value for customers.

## Inputs and Deflators

### Personal Social Services

- 10.53** The sources and composition of Personal Social Services inputs and deflators rely mainly on local authority information. More detailed work is required to scope a full work programme but it is likely that there will be a similar need to improve cost deflators used by ONS. It would also be helpful to improve the timeliness of cost information provided by local authorities.
- 10.54** The independent sector (for-profit and non-profit) has a substantial role in the delivery of adult social care services, funded through local authority contracts. Moreover, care homes, home care agencies and other providers generally provide equivalent or similar services to both publicly-funded and privately-funded clients. It may be important to analyse information about the independent residential care and home care market to see whether this helps understand local authority expenditure on goods and services in this area.

### Social Security

- 10.55** *Blue Book 2003* showed a marked increase in Social Protection spending in 2002, the first full year of the newly formed DWP. It seems likely that this is partly attributable to misclassification errors following the setting up of the new department. Misclassification errors on government inputs, even if they are associated with offsetting errors elsewhere in the government sector, could affect published GDP if different output measurement methods are applied in the affected areas. Preparation of the 2004 *Blue Book* has, we understand, allowed some progress on this front.
- 10.56** To clear up these matters thoroughly, a comprehensive examination of the inputs data relating to the department is required. This should ensure continuity and examine the split between capital and current expenditure, and current expenditure that builds capacity for the future.
- 10.57** The principle applies more widely, to any other areas affected by machinery of government changes: it is clear that careful review is needed to ensure that inputs and relevant outputs are transferred together, and take account of new developments. Accurate tracking of expenditure switches within government and regular review of outputs are needed to ensure well-founded and comprehensive input and output measures in the National Accounts.

## Triangulation and Productivity Measurement

- 10.58** Triangulation will be particularly important for social services, given the difficulty in measuring quality changes for services to individuals. Other evidence on benefits to clients, from research or user surveys, may tell a clearer story than a simple productivity measure derived from the ratio of outputs to inputs. We will work further with the departments to consider what evidence might be available, with the aim of publishing a productivity article in the future.

## Summary of Work in Progress and Future Directions

**10.59** During the rest of our review, we will:

- a) work with DH and ONS to develop an improved social services cost-weighted activity series with more complete coverage of adult social care;
- b) work with DfES to ensure that similar improvements are made in the output series for children's social services;
- c) work with DH to consider the focus of academic research on the measurement of output and productivity in adult social care most relevant for the National Accounts;
- d) work with DfES on ways of using quality or outcome measures for children's social services to improve the current activity measures;
- e) work with those in Scotland, Wales and Northern Ireland responsible for adult and children's social services to assess what data and analysis might be available, to avoid using England as a proxy for the United Kingdom as a whole;
- f) identify any changes needed in deflators used by ONS for social services expenditure, with improved sources and techniques if possible;
- g) consider what information about the market sector in residential and home care would help our understanding of publicly provided care;
- h) work with DWP to develop an output indicator for social security that takes account of all relevant information on benefit and pension claim handling, including quality indicators; and also captures broader policy objectives such as welfare to work and the provision of direct support to people with disabilities or to other specific groups of people; and
- i) work with DWP and ONS to ensure DWP inputs (and those of other departments affected by machinery of government changes) are reviewed thoroughly, and that misallocation issues are resolved.

# 11 Summary of Main Conclusions

**11.1** This chapter summarises our interim conclusions (discussed in Chapters 5 and 6), including the principles we are applying in developing measures of the change over time in government output at constant prices. It also summarises the work in progress in Health, Education, Public Order and Safety and Social Protection.

## Interim Conclusion (1) Regarding Principles:

The direct measurement of the output from government spending, and the measurement of inputs and productivity, should be based on a set of principles, within the framework set by international guidelines.

These principles (discussed in Chapter 5) are:

**Principle A:** the objective should be to measure output from government spending in terms of its incremental contribution to individual or collective welfare, in the same way as market output.

**Principle B:** the procedure of defining direct output indicators within a government function should start by seeking to identify the services provided by government to households and firms, and attempts made to find data to reflect these services as comprehensively as possible (rather than working back from available indicators). Where, initially, it is necessary to apply an indicator from another service, this should be explicit. The coverage of indicators within a function should be reassessed on a regular basis.

**Principle C:** formal criteria should be set in place for the further extension of direct output measures. 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 allowance for quality change, (iii) the effects of its introduction have been tested service by service, and (iv) the context in which they will be published has been fully assessed, in particular the implied productivity estimate.

**Principle D:** 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 E:** the measurement of inputs should be as comprehensive as possible and, in particular, should include capital services; consideration should be given to the split between current and capital spending.

**Principle F:** criteria should be established for the quality of 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 G:** independent corroborative evidence should be sought on government productivity, as part of a process of ‘triangulation’, taking account of the timing of inputs and outputs.

**Principle H:** value should be seen as adjusted for quality; for each service, explicit consideration should be given to the incorporation of quality change as an element of value added; for each spending function, consideration should be given to the extent to which quality change is captured by the changing activity mix, and to the way in which output measures for government should be adjusted for increased real value in an economy with rising real GDP.

### **Interim Conclusion (2) Regarding Inputs:**

The consideration of the measurement of inputs, and of the price deflators applied, has turned out to be an important element of the review; and will form a significant part of our work in the second stage.

### **Interim Conclusion (3) Regarding Extension of Direct Output Measures:**

The overall thrust of the Commission Decision is welcome and we believe that ONS, like other statistical offices, should work vigorously on this agenda. Implementation decisions, however, should be taken with due caution against a clear knowledge of the implications and against clear criteria.

### **Interim Conclusion (4) Regarding Input Data:**

The conclusions are:

- a) ONS, the review team and the Treasury should work together to review the data supply and processing chain for central government estimates, to feed into the Treasury review already underway;
- b) ONS and the review team and ODPM should agree an action plan to address data timeliness, breakdown, and periodicity of local authority data;
- c) ONS, the review team, with the assistance of the Treasury and ODPM should carry out a study to establish whether ONS should replace the capital consumption estimates modelled by ONS with the capital consumption estimates supplied by the Treasury and ODPM; and
- d) ONS and review team, together with relevant departments, should carry out a review of pay and price deflators.

### Interim Conclusion (5) Regarding Presentation:

Greater consideration should be given to the presentation of National Accounts measures of government output and productivity, bringing out the purposes for which they should and should not be used. National Accounts should be coherent with, but are different from, PSA targets and micro-measures of public sector efficiency; explicit reference should be made to the margins of error surrounding estimates.

### Interim Conclusion (6) Regarding Documentation:

ONS and the review team should work closely with relevant government departments to ensure that documentation is produced that fully informs all parties, including departments and the public, about the processes involved in compiling the government expenditure component of the National Accounts.

### Interim Conclusion (7) Regarding Resources:

If the government wishes to have reliable estimates of government output and productivity, then the statistical resources have to be supplied.

- 11.2** These conclusions and principles are reflected in the analysis of issues for the main areas of general government expenditure in Chapters 7–10, and the summary below of work in progress for the rest of the review.

### Health

- 11.3** Our main concerns are that current measures of health care used in the National Accounts:
- a) do not include quality measures, which are an essential part of output (as distinct from activity);
  - b) assume that health care activity growth in England is a proxy for all UK activity, which may not be correct; and
  - c) rely on the General Household Survey to estimate outputs of general practitioners and practice staff.
- 11.4** During the remaining period of our review, we will undertake the following work:
- a) We will work with DH on sensitivity testing, using any sample or local data which may be available, to consider which health care activities not currently measured nationally should be particular priorities for better measures.
  - b) We will work with ONS experts in the General Household Survey to consider what the error rates may be in the estimated GP consultation rates and GDP estimates based on this source. If this confirms a need for improvement, we will review with DH whether better methods can be found to measure changes in services provided by GPs, taking account of the new GP contract and other structural changes. One possibility may be the use of a long-term database of information about a subset of GP practices such as the Royal College of General Practitioners spotter practices or

General Practice Research Database, though the practices involved may not be a representative sample.

- c) We will consider further how to make use of an approach which uses common cost weights for treatments which are substitutes, especially if there are differences in quality of outcome for the patient and if all costs and benefits of a complete course of treatment can be identified. This might be part of a triangulation approach.
- d) We will continue discussions with the York/NIESR research team who are advising DH on measuring NHS outputs and productivity, so we can take account of their work (which continues until late 2005) when we comment in our final report on the most promising directions for incorporating quality measures into Health outputs as used in the National Accounts.
- e) We will work with ONS, DH and the health departments in Wales, Scotland and Northern Ireland to see how far relevant data from their respective health systems can be incorporated into the UK National Accounts, using the techniques adopted for England in 2004 or adapting them as appropriate, given differences in systems, under the principles set out in this report.
- f) We will work with ONS and DH to improve techniques to estimate labour and capital components of input estimates at constant prices, including more accurate ways of reflecting changes in skill mix and relative earnings.

**11.5** For the longer term, we see benefits in further work to explore new measures of NHS work which combine separate activities into ‘care pathways’, to give the clearest picture of value added by the NHS. We consider it important that developments in NHS information systems, while intended primarily to support patient care, should also support better future analysis of health care outputs.

## Education

**11.6** During the rest of the review, we will work with ONS, DfES and the education departments of the devolved administrations to undertake the following work.

- a) We will examine whether a volume measure can be produced that is closer to pupil hours taught than the current measure of full time equivalent pupils. This would look at length of school days and absence rates.
- b) We will explore further two approaches to the measurement of quality in education as part of value added by schools: the measure of pupil progress based on attainment through examination results; and on a measure of the quality of teaching, or overall school performance, derived from Ofsted inspections. If further work suggests both are viable, they could be weighted together in a value-added measure: if so, we would need to consider carefully what their relative weights should be.
- c) We will seek additional measures of school examination attainment post-16, since this has been an area of increased expenditure not matched by current proposed measures of attainment.

- d) We will consider whether evidence of the value of education which reflects the wider economy, such as net present value of future earnings and the benefits of child care provided by schools, could be used in the National Accounts.
- e) We will examine further the new output measure introduced for health professional education and consider what improvements might be made, including ways of incorporating quality change.
- f) We will consider whether there are other areas where government expenditure on education is not at present classified as such, and, if so, what specific output measures should be used.
- g) We will improve the analysis of input data by identifying current inputs that are intended to build capacity for the future or can only deliver outputs after a lag, and improvement to pay and price indices.

**11.7** We are conscious that our work is not covering the higher or further education sectors at all, since these are by definition outside the government sector of the National Accounts and so outside our terms of reference. In the longer-term, given the importance of education to the economy, we feel it would be worthwhile for ONS to look at whether its current measures are capturing the added value of this sector, perhaps as part of the suggested satellite account on formation of human capital.

### **Public Order and Safety**

**11.8** Our main concern is to ensure that the measures of a collective service reflect the added value to society, rather than simply measuring volumes of activities. This may best be achieved by looking at overall outcomes rather than attempting to ascribe quality to individual activities. For the civil courts, however, an approach to valuing individual activities for individual end users is more applicable.

**11.9** During the rest of the review we will:

- a) work closely with the Home Office, DCA, ODPM and other relevant departments and agencies to explore how much of government expenditure on Public Order and Safety can be tied to measurable output, including methods of building quality into the measurement;
- b) consider a triangulation approach using independent productivity measures or measures of the effectiveness of the whole system, where expenditure does not match activities whose value added can be measured;
- c) work with the Home Office and other departments on the alternative approaches to measuring output discussed above;
- d) work with DCA to improve measures of the civil courts' output, ensuring all output is captured, that expenditure data for the construction of weights is improved and a robust measure of quality change is developed;

- e) explore whether a number of different measures of fire service productivity could be used to assess the overall value added by the service, with a triangulation approach to be considered as a possible improvement on current activity measures;
- f) work with ONS and the relevant departments to ensure expenditure data are complete for each function and are correctly assigned to relevant outputs;
- g) identify aspects of expenditure, like staff training that should not be expected to lead immediately to extra activity and consider how to take account of this expenditure in developing and interpreting productivity measures;
- h) work with ONS and departments to ensure satisfactory deflators are available for each expenditure area; and
- i) work with devolved administrations towards using direct data, having regard to the separate systems in place in the various countries, to underpin the measure of UK output.

### Social Protection

**11.10** During the rest of our review, we will:

- a) work with DH and ONS to develop an improved social services cost-weighted activity series with more complete coverage of adult social care;
- b) work with DfES to ensure that similar improvements are made in the output series for children's social services;
- c) work with DH to consider the focus of academic research on the measurement of output and productivity in adult social care most relevant for the National Accounts;
- d) work with DfES on ways of using quality or outcome measures for children's social services to improve the current activity measures;
- e) work with those in Scotland, Wales and Northern Ireland responsible for adult and children's social services to assess what data and analysis might be available, to avoid using England as a proxy for the United Kingdom as a whole;
- f) identify any changes needed in deflators used by ONS for social services expenditure, with improved sources and techniques if possible;
- g) consider what information about the market sector in residential and home care would help our understanding of publicly provided care;
- h) work with DWP to develop an output indicator for social security that takes account of all relevant information on benefit and pension claim handling, including quality indicators; and also captures broader policy objectives such as welfare to work and the provision of direct support to people with disabilities or to other specific groups of people; and

- i) work with DWP and ONS to ensure DWP inputs (and those of other departments affected by machinery of government changes) are reviewed thoroughly, and that misallocation issues are resolved.

## Appendix A: UK Consultation

During the first few months of the report, we have consulted with, or given presentations, to the following:

Advanced Institute of Management Research  
 The Audit Commission  
 Bank of England and the Monetary Policy Committee  
 Institute of Fiscal Studies  
 National Audit Office  
 Ofsted  
 Prime Minister's Delivery Unit  
 The Royal Statistical Society  
 Statistics Commission  
 Centre for Health Economics, University of York  
 Northern Ireland Executive  
 Scottish Executive  
 Welsh Assembly

We received assistance from the Coordination group which is made up as follows:

<b>Government Dept</b>	<b>Position</b>	<b>Position</b>
Office for National Statistics	Len Cook	National Statistician - Chair
Office for National Statistics	Colin Mowl	Executive Director, Macroeconomics & Labour Market
Bank of England	Charles Bean	Chief Economist
Department for Education & Skills	Paul Johnson	Head of Analytical Services
Department for Education & Skills	Stephen Kershaw	Director of Finance
Department of Health	Professor Barry McCormick	Chief Economist
Department of Health	Richard Douglas	Director of Finance & Investment
Department of Health	John Fox	Director of Statistics
Department of Health	Martin Campbell	Finance and Investment Directorate
Department of Work & Pensions	Nick Dyson	Director of Information & Analysis
Department of Work & Pensions	Sue Rice	Senior Manager, Analytical Support
HM Treasury	Sir Nicholas Stern	Chief Economist
HM Treasury	Jonathan Stephens	Director of Public Services
Home Office	Professor Paul Wiles	Director, Research, Development & Statistics
Home Office	Dr Stephen Almond	Economic Advisor
Home Office	William Nye	Director of Performance and Finance
Office of the Deputy Prime Minister	Michael Kell	Director of Analysis and Research
Office of the Deputy Prime Minister	Meg Green	Divisional Manager of Statistics

Many other civil servants from these departments helped our work.

## Appendix B: International Consultation

International Body	Name	Position
Australian Bureau of Statistics	Rob Edwards	Deputy Australian Statistician – Economic Statistics
	Peter Harper	First Assistant Statistician – Economic Statistics
Brookings Institution	Jack Triplett	Research Fellow
Bureau for Economic Analysis (USA)	Dr Barbara Fraumeni	Chief Economist
Eurostat	Michel Vanden Abeele	Director General
	Brian Newson	Acting Director - Economic and Monetary Statistics
	Paul Konijn	Unit C1, European System of Accounts
Government Institute for Economic Research – Finland	Kalevi Luoma	Research Chief
National Institute of Statistics (Italy)	Luigi Biggeri	President
	Alfonsina Caricchia	Head of Department of Integration and Technical Standards
	Daniela Collesi	Senior Researcher
	Silvia Zannoni	Researcher
Office of Economic and Statistical Research – Queensland, Australia	Massimo Anzalone	Researcher
	Peter Crossman	Assistant Under Treasurer and Government Statistician
Organisation for Economic Co-operation and Development (OECD)	Enrico Giovanninni	Head of Statistics Division
	Francois Lequiller	Head of National Accounts
	Paul Schreyer	Statistics Division
	Dave Turner	UK Country Desk
	Jens Lunsgaard	UK Country Desk
	Manfred Huber	Department of Health Data and Health Accounting
	Jon Blondal	Budgeting and Marketing Division
Michael Ruffner	Budgeting and Marketing Division	
Statistics Finland	Raimo Nurminen	Senior Statistician – National Accounts
	Olli Seppanen	Researcher – National Accounts
Statistics Netherlands	Peter van der Pen	Head of National Accounts
	Sake de Boer	Government Output Measurement
	Foske Kleima	Health and Education

<b>International Body</b>	<b>Name</b>	<b>Position</b>
Statistics Norway	Svein Longva	Director General
	Olav Ljones	Assistant Director General
	Knut Sorensen	Research Division for National Accounts
	Anne-Britt Svinnset	Head of Division- Public Finance
	Tor Skoglund	Researcher - National Accounts
	Ann Lisbet Brathaug	Head of Division - National Accounts
Statistics Sweden	Per Ericsson	Head of National Accounts
	Brigit Magnusson	Deputy Head of National Accounts
	Jenny Davidsson	National Accounts
	Ann-Louis Paivinen	National Accounts
	Anna Widenfalk	National Financial Management Authority

# Appendix C: International Guidance

---

- C1** This appendix gathers the international guidance on compilation of General Government Final Consumption (GGFC) for national accounts, particularly focusing on text which raises key issues. Some of the guidelines included are general to national accounts whereas some are specific to the general government sector. The implications of the ongoing review of the System of national accounts for GGFC are also explored. As productivity is not a national accounts concept, the guidance for this area is distinguished from national accounts guidance and OECD guidance on estimation of productivity outside of the national accounts framework has also been included.

The appendix covers the following areas:

- a) Sources of International Guidelines
- b) Timing of Recording in National Accounts
- c) General Government Sector and Non-Market Producers
- d) Classification of Government by Function
- e) GGFC Expenditure and Actual Final Consumption
- f) General Government as a Producer
- g) Valuation of Government Output
- h) Measuring the Volume of Government Output
- i) Capital Measures and Definition
- j) OECD Guidance on Productivity
- k) ESA Guidance on Satellite Accounts

Table C1 Sources of international guidance used as reference material for the Atkinson Review of Measurement of Government Output

Title	Organisation/s	Status	Type of guidance
System of National Accounts (1993)	UN, OECD, World Bank, Commission of the European Communities and the IMF. Document prepared by Inter-secretariat Working Group on National Accounts and approved by UN Statistics Commission.	Recommended but not mandatory - recognises need for flexibility. Currently under review for an SNA 2008 edition.	High-level guidance only
European System of Accounts (1993)	Eurostat	An EU version of SNA but with a legal basis to ensure 'strict application' to ensure 'harmonised and reliable statistics on which to base decisions'. Inventories of methods carried out country by country.	Broadly consistent with SNA (93) guidance with more clarification.
Handbook on prices and volume in National Accounts	Eurostat	Driven by the needs of the 1997 Growth and Stability Pact. SNA 93/ESA 95 'not sufficient to guarantee harmonised price and volume data'. Introduces an A/B/C score for methods.	Expansion of ESA (93) guidance for government output distinguishing activities, outputs and outcomes and with examples of each.
Measuring Productivity	OECD	One of its objectives: to improve international harmonisation: although there is no strong prescriptive element in the manual, it contains indications about desirable properties of productivity measures.	'Comprehensive guide to the various productivity measures aimed at statisticians, researchers and analysts involved in constructing industry-level productivity indicators'.
United Nations website	UN	Guidance for Classification by Function of Government	<a href="http://unstats.un.org/unsd/cr/registry/default.asp">http://unstats.un.org/unsd/cr/registry/default.asp</a>

- C2** The following sections are divided by guidance topics, with different sources of guidance where applicable and an outline of the UK practice in relation to the guidance given. The text shown is not exhaustive but has been selected to illustrate the key issues.

### Timing of Recording in National Accounts

#### SNA guidance

- C3** ‘The System recommends recording on an accrual basis throughout.’ (paragraph 3.91)

‘Accrual accounting records flows at the time economic value is created, transformed, exchanged, transferred or extinguished. This means that flows which imply a change of ownership are entered when the ownership passes, services are recorded when provided, output at the time products are created and intermediate consumption when the materials and supplies are being used’ (paragraph 3.94)

‘Furthermore, some transactions, in particular government units, do not keep records of purchases on an accruals basis. In these cases, the rules of consistency in the System require that efforts should be undertaken to correct basic statistics for major deviations and flaws.’ (paragraph 3.96)

#### ESA guidance

- C4** ‘However, in some cases it is necessary to show flexibility as regards time of recording. This applies in particular to taxes and other flows concerning general government, which are often recorded on a cash basis in government accounts. It is sometimes difficult to carry out an exact transformation of these flows from cash basis to accrual basis. In these cases, it might therefore be better to use approximations ... Consequently, transactions may be recorded at different times by the transactors involved. These discrepancies must be eliminated by adjustments.’ (paragraph 1.57)

‘Output is to be recorded and valued when it is generated by the production process.’ (paragraph 3.46)

ESA 95 was modified in two legal acts:

- C5** i) The following text is in *Regulation (EC) No 2516/2000* published Nov 2000: ‘General principles: The impact on net lending/borrowing of general government of taxes and social contributions recorded in the system shall not include amounts unlikely to be collected...’ and ‘Taxes and social contributions recorded in the accounts may be derived from two sources: amounts evidenced by assessments and declarations and cash receipts ... if assessments and declarations are used, the amounts shall be adjusted by a coefficient reflecting assessed and declared amounts never collected ... if cash receipts are used, they shall be time-adjusted so that the cash is attributed when the activity took place to generate the tax liability ...’
- ii) The following text is in *Regulation (EC) No 995/2001* published May 2001 ‘... taxes and social contributions payable to the general government can either be recorded net of the part unlikely to be collected or, if this part is included, it should be neutralised in the same accounting period by a capital transfer from the general government to the relevant sectors.’

## General Government Sector and Non-Market Producers

### SNA guidance

- C6** ‘General government: institutional units which, in addition to fulfilling their political responsibilities and their role of economic regulation, produce principally non-market services (possibly goods) for individual or collective consumption and redistribute income and wealth’ (paragraph 2.20)

‘The general government sector consists mainly of central, state and local government units together with social security funds imposed and controlled by those units. In addition, it includes non-profit institutions engaged in non-market production that are controlled and mainly financed by government units or social security funds.’ (paragraph 4.9)

### ESA guidance

- C7** ‘for the purposes of the system, the institutional units are grouped together into five mutually exclusive institutional sectors composed of the following types of units: (paragraph 1.28)

- a) non-financial corporations;
- b) financial corporations;
- c) general government;
- d) households;
- e) non-profit institutions serving households ...’

- C8** ‘The sector general government (S.13) includes all institutional units which are other non-market producers ... whose output is intended for individual and collective consumption, and mainly financed by compulsory payments made by units belonging to other sectors, and/or all institutional units principally engaged in the redistribution of national income and wealth.

- C9** ‘The institutional units included in sector S.13 are the following:

- a) ‘general government entities (excluding public producers organised as public corporations or, by virtue of special legislation, recognised as independent legal entities, or quasi-corporations, when any of these are classified in the non-financial or financial sectors) which administer or finance a group of activities, principally providing non-market goods and services, intended for the benefit of the community;
- b) ‘non-profit institutions recognised as legal entities which are other non-market producers and which are controlled and mainly financed by general government;...

‘The general government sector is divided into four sub-sectors: (paragraph 2.70)

- a) central government;
- b) state government;
- c) local government;
- d) social security funds; (paragraphs 2.68-2.70)

**C10** ‘Definition: Other non-market producers are local kind-of-activity units or institutional units whose major part of output is provided free or at not economically significant prices.’ (paragraph 3.26)

‘If less than 50 per cent of the production costs are covered by sales, the institutional units is an other non-market producer and classified to the sector Non Profit Institutions serving Households (NPISH). But other non-market non profit institutions that are controlled and mainly financed by general government are classified to the general government sector.

**C11** ‘In distinguishing market and other non-market producers by mean of the 50 per cent criterion, sales and costs are defined as follows:

- a) sales cover the sales excluding taxes on products but including all payments made by general government ... and granted to any kind of producer in this type of activity, i.e. all payments linked to the volume or value of output are included, but payments to cover an overall deficit are excluded ...
- b) production costs are the sum of intermediate consumption, compensation of employees, consumption of fixed capital and other taxes on production. For this criterion other subsidies on production are not deducted. To ensure consistency of the concepts sales and production costs when applying the 50 per cent criterion, the production costs should exclude all costs made for own-account capital formation.’

**C12** The 50 per cent criterion should be applied by looking over a range of years: only if the criterion holds for several years or holds for the present year and is expected to hold for the near future, it should be applied strictly. Minor fluctuations in the size of sales from one year to another do not necessitate a reclassification of institutional units ...’ (paragraphs 3.32-3.33)

## Classification of Government by Function

**C13** SNA guidance (on collective and individual services)

### Individual services

‘Individual goods and services are essentially ‘private’, as distinct from ‘public’ goods. They have the following characteristics: (paragraph 9.81)

- a) It must be possible to observe and record the acquisition of the good or service by an individual household or member thereof and also the time at which it took place;
- b) The household must have agreed to the provision of the good or service and take whatever action is necessary to make it possible – for example, by attending a school or clinic;
- c) The good or service must be such that its acquisition by one household or person, or possibly by a small restricted group of persons, precludes its acquisition by other households or person ...

‘From a welfare point of view, the important characteristic of an individual good or service is that its acquisition by one household, person or group of persons brings no (or very little) benefit to the rest of the community ... (paragraph 9.82)

### Collective services

**C14** ‘Most goods can be privately owned and are individual in the sense used here. On the other hand, certain types of services can be provided collectively to the community as a whole. The characteristics of these services may be summarised as follows:

- a) Collective services can be delivered simultaneously to every member of the community or of particular sections of the community, such as those in a particular region of a locality;
- b) The use of such services is usually passive and does not require the explicit agreement or active participation of all the individuals concerned;
- c) The provision of a collective service to one individual does not reduce the amount available to others in the same community or section of the community. There is no rivalry in acquisition.’ (paragraphs 9.81-9.83)

‘Expenditures incurred by governments at a national level in connection with individual services such as health and education are to be treated as collective when they are concerned with the formulation and administration of government policy, the setting and enforcing of public standards, the regulation, licensing or supervision of producers, etc ... on the other hand, any overhead expenses connected with the administration or functioning of a group of hospitals, schools, colleges or similar institutions are to be included in individual expenditures ...’ (paragraph 9.86)

### SNA guidance on classification of function of government

**C15** ‘... it may not be possible to classify transactions and, as an approximation, the units of classification may have to be agencies, offices, bureaux or project units within government departments ... it may happen of course that the smallest units that can be identified still perform two or more classification of the functions of government functions; in such cases it will be usually be best to make an approximate division of the unit's outlays among the different functions performed rather than to allocate them all to that which is judged the largest.’ (paragraph 18.10)

Table C2 **Classification of the functions of government (COFOG) used for compilation of the expenditure measure of GDP. The COFOG classifications are guidelines provided in the System of National Accounts 1993 (UN,OECD, IMF, CEC and World Bank).**

	Divisions	Sub-divisions
Collective services	01 General Public Services	01.1 Executive and legislative organs, financial and fiscal affairs, external affairs 01.2 Foreign economic aid 01.3 General services 01.4 Basic research 01.5 R&D General public services 01.6 General public services n.e.c. 01.7 Public debt transactions 01.8 Transfers of a general character between different levels of government
	02 Defence	02.1 Military defence 02.2 Civil defence 02.3 Foreign military aid 02.4 R&D defence 02.5 Defence n.e.c.
	03 Public Order and Safety	03.1 Police services 03.2 Fire-protection services 03.3 Law courts 03.4 Prisons 03.5 R&D Public order and safety
	04 Economic Affairs	04.1 General economic, commercial and labour affairs 04.2 Agriculture, forestry, fishing and hunting 04.3 Fuel and energy 04.4 Mining, manufacturing and construction 04.5 Transport 04.6 Communication 04.7 Other sectors 04.8 R&D Economic Affairs 04.9 Economic affairs n.e.c.
	05 Environmental Protection	05.1 Waste management 05.2 Waste water management 05.3 Pollution abatement 05.4 Protection of biodiversity and landscape 05.5 R&D Environmental protection 05.6 Environmental protection n.e.c.
Individual services	06 Housing and Community Amenities	06.1 Housing development 06.2 Community development 06.3 Water supply 06.4 Street lighting 06.5 R&D Housing and amenities 06.6 Housing and community amenities n.e.c.

	<b>Divisions</b>	<b>Sub-divisions</b>
Individual services	07 Health	07.1 Medical products, appliances and equipment 07.2 Out-patient services 07.3 Hospital services 07.4 Public health services 07.5 R&D health 07.6 Health n.e.c.
	08 Recreation, Culture and Religion	08.1 Recreational services 08.2 Cultural services 08.3 Broadcasting and publishing services 08.4 Religious and other community services 08.5 R&D Recreation, culture and religion 08.6 Recreation, culture and religion n.e.c.
	09 Education	09.1 Pre-primary and primary education 09.2 Secondary education 09.3 Post-secondary non-tertiary education 09.4 Tertiary education 09.5 Subsidiary services to education 09.6 R&D Education 09.7 Education n.e.c.
	10 Social Protection	10.1 Sickness and disability 10.2 Old age 10.3 Survivors 10.4 Family and children 10.5 Unemployment 10.6 Housing 10.7 Social exclusion n.e.c. 10.8 R&D Social protection 10.9 Social protection n.e.c.

## GGFC Expenditure and Actual Final Consumption

### SNA guidance

- C16** ‘Final consumption expenditure covers transactions on final consumption of goods and services for which a sector is the ultimate bearer of the expense. Government and NPISH produce non-market goods and services in their production account, where intermediate consumption or compensation of employees are recorded as uses. Final consumption expenditure of these producers relates to the value of their output of non-market goods and services, less their receipts from the sale of non-market goods and services at prices which are not economically significant. However, it also covers services that are purchased by government or NPISHs for ultimate transfer, without transformation, to households.’

‘The use of adjusted disposable income account (...) records adjusted disposable income as resources and actual final consumption as uses... Actual final consumption of households covers goods and services which are effectively available for individual consumption by households, regardless of whether the ultimate bearer of the expense is government, NPISHs or households themselves. Consequently, actual final consumption of government refers only to collective consumption, whereas NPISHs, whose final consumption expenditure is deemed to be in total individual, have no actual final consumption.’ (paragraphs 2.127-2.128)

### ESA guidance

- C17** ‘The use of adjusted disposable income account includes the concept of actual final consumption, which corresponds to the value of the goods and services actually at the disposal of households for final consumption, even if their acquisition is financed by general government or non-profit institutions serving households.’

‘Consequently, the actual final consumption of general government corresponds only to collective final consumption. Since final consumption expenditure by non-profit institutions serving households is regarded as entirely individual, their actual final consumption is zero.’

‘At the level of total economy, final consumption expenditure and actual final consumption expenditure are equal ...’ (paragraphs 8.40-8.41)

- C18** ‘Final consumption expenditures by general government or NPISHs are equal to the sum of their output, plus the expenditures on products supplied to households via market producers (i.e. Social transfers in kind) minus the payments by other units minus own-account capital formation.’ (paragraph 3.96)

## General Government as a Producer

### SNA guidance

- C19** ‘Establishments owned by government or NPISHs commonly provide education, health or other services to individual households without charge or at prices that are not economically significant. The costs of providing these services are incurred by the government or NPISHs, and the values are recorded as internal transactions: that is, as final expenditures by governments or NPISHs on outputs produced by establishments they own themselves. ( ... the acquisition of these services by households is recorded separately under social transfers in kind, another form of non-monetary transactions that take place between the government units or NPISHs and the households in question.)’ (paragraph 3.47)

- C20** ‘Non-market producers providing final goods or services – such as public administration, defence, health and education – should be partitioned into establishments using the activity classification given in divisions 75, 80, 85 and 90 of the ISIC ...’ (paragraph 5.38)

## Valuation of Government Output

### SNA guidance

- C21** ‘In contrast to output produced for own consumption or own gross capital formation by market producers, there are usually no suitable markets whose prices can be used to value government non-market output. By convention, therefore, such output is valued by its production costs.’ (paragraph 4.110)

- C22** ‘In the System, the intermediate inputs are recorded and valued at the time they enter the production process, while outputs are recorded and valued as they emerge from the process ... The increase between the value of intermediate inputs and the value of outputs is the gross value added against which must be charged the consumption of fixed capital, taxes on production (less subsidies) and compensation of employees. The positive or negative balance remaining is the net operating surplus or mixed income. The definition measurement and valuation of outputs and inputs is, therefore, fundamental to the System.

‘Output therefore consists only of those goods or services that are produced within an establishment that become available for use outside that establishment ...

‘For simplicity, the output of most goods or services is usually recorded when their production is completed ...’ (paragraphs 6.37-6.38)

- C23** ‘There are no markets for collective services such as public administration or defence, but even in the case of non-market education, health or other services provided to individual households, suitable prices may not be available. It is not uncommon for similar kinds of services to be produced on a market basis and sold alongside the non-market services but there are usually important differences between types and quality of services provided. In most cases it is not possible to find enough market services that are sufficiently similar to the corresponding non-market services to enable their prices to be used to value the latter, especially when the non-market services are produced in very large quantities.

For these reasons, and also to ensure that the various non-market services produced by government and NPISHs are valued consistently with each other, they are all valued in the System by the sum of costs incurred in their production: that is, as the sum of:

- intermediate consumption;
- compensation of employees;
- consumption of fixed capital; and
- other taxes, less subsidies, on production.

The net operating surplus on the production of non-market goods or services produced by government units and NPISHs is assumed always to be zero.’ (paragraphs 6.90-6.91)

## Measuring the Volume of Government Output

### SNA guidance

- C24** ‘In principle, volume indices may always be compiled directly by calculating a weighted average of the quantity relatives for the various goods and services produced as outputs using the values of these goods and services as weights. Exactly the same method may be applied even when the output values have been estimated on the basis of their costs of production.

Of course, the calculation of quantity relatives for the outputs of many kinds of non-market services, especially collective services, presents problems. In the case of health and education services provided as social transfers to households, however, the problems are much less, both conceptually and in practice, than for collective services such as public administration or defence. The objective is to measure the quantities of the services actually delivered to households. These should not be confused with the benefits or utility derived from those services ...’ (paragraphs 16.134-16.135)

- C25** ‘Measuring changes in the volume of collective services is distinctly more difficult, however, as it is not possible to observe and record the delivery of such services. Many collective services are preventative in nature; protecting households or other institutional units from acts of violence, including acts of war, or protecting them from other hazards, such as road accidents, pollution, fire, theft or avoidable diseases. It is difficult to measure the output of preventive services, and this is an area in which further research is needed. In practice, it may not be feasible to avoid using changes in the volumes of inputs into such services as proxies for changes in volumes of outputs ...

‘When it is not possible to avoid using an input measure as a proxy for an output measure, the input measure should be a comprehensive one and not confined to labour inputs ... These volume measures can, of course, also be derived by deflating the current values by suitably weighted wage rate, price or tax rate indices.’ (paragraphs 16.139-16.140)

### ESA guidance

- C26** ‘The establishment of a comprehensive system of price and volume indices covering all supply and uses of goods and services encounters a particularly difficulty when measuring the output of non-market services. These services differ from market services in that they are not sold at a market price and their value at current prices is calculated by convention as the sum of the costs incurred. These costs are intermediate consumption, compensation of employees, other taxes less subsidies on production and consumption of fixed capital.

‘In the absence of a unit market price, the change in the “unit cost” of a non-market service can be considered as an approximation of the change in the price. If non-market services are consumed on an individual basis, it is in principle possible to estimate quantities which are homogeneous and which reflect the utilisation of these services and apply the unit costs of a base year to obtain data in constant prices. By such type of output-measurement it will be possible to analyse changes in productivity for individual non-market services. For collective services it is generally not possible to establish unit costs and quantities reflecting their utilisation. If attempts are made to account for changes in productivity for collective services by indirect methods, users should be made aware of this.

‘In the context of the economic accounts, it is of prime importance to adopt the principle that the production and consumption of non-market services, must be defined in terms of the actual flows of these goods and services and not in terms of the final results obtained from their use. As these results depend on several other factors as well, it is not possible to measure, for example, the volume of teaching services by the rise in the level of education, or the volume of health services by the improvement in the health of the population.’ (paragraphs 10.24- 10.26)

**C27** ‘For certain market and non-market service industries, such as finance, business services, education or defence, it may not be possible to obtain satisfactory estimates of price or volume change for output. In these cases the movements of value added at constant prices can be estimated by means of changes in compensation of employees at constant wage rates and consumption of fixed capital at constant prices. Compilers of data may be forced to adopt such expedients, even when there is no good reason to assume that labour productivity remains unchanged in the short or long term.’ (paragraph 10.29)

**C28** ‘In the case of services provided to individuals, changes in the volume of their output and consumption should in principle be measured on the basis of the *use [sic]* which is made of these services; this will avoid using different criteria for the same service depending on whether they are market or non-market. Of course, any change in quality must be treated as a change in volume; but this applies as much to market services as to non-market services provided to individuals.

‘The pure collective services are produced by general government for the benefit of the entire population. In fact, they cover a vast range of activities such as general public services, national defence, foreign affairs, justice and the police, town planning and the environment, economic policy, etc. Since these services are consumed collectively, indirectly and continuously, the volume of their output cannot be measured by the extent to which they are utilised’. (paragraphs 10.42- 10.43)

### Eurostat Handbook guidance

**C29** ‘Non-market output can only be produced by non-market producers ... that may or may not also produce market outputs. The total value of output of a non-market producer is defined by convention as the total costs of production (i.e. the operating surplus is assumed to be zero). In the case of a local KAU [kind of activity unit] with secondary market output, non-market output is defined as a residual item, i.e. as the difference between the total costs of production minus the revenues from market output.

‘It is important to note that this valuation principle (calculating current price output as costs) is applied to the producer rather than the product. Non-market producers are either public producers or non-profit institutions, classified in the sectors government or NPISH respectively.

Non-market output can be sub-divided into two types of output:

- individual goods and services: those that are consumed by individual households; and
- collective services: those that are provided simultaneously to the society as a whole (by definition, goods can not be collective).

‘Examples of individual products are education, health, social security, recreation services and cultural services. Examples of collective services are general public administration, defence, police services and research and development.’ (section 3.1.2)

#### ‘Input, activity, output and outcome

‘The following criteria can be formulated for the appropriate use of output indicators:

- they should cover all the services produced by the producer that are provided to external users, and only those; activities that are in fact ancillary to the main output should not be counted;
- they should be weighted by the costs of each type of output in the base year;
- they should be defined as detailed as possible;
- they should be quality-adjusted.’

**C30** ‘The problem of measuring prices and volumes for non-market output arises from the fact that by definition no market prices exist. For that reason, the value of output at current prices is defined as the sum of costs minus revenues as noted above. Without prices for the output, there are only two options for constant price measurement: deflating inputs and direct volume measurement.

‘Current practice for constant prices is mostly based on deflating inputs. This implies assuming that the change in volume of inputs is representative for the change in the volume of output. However, it is not at all certain that more or better inputs lead automatically to more or better output. Using this assumption makes it impossible to analyse change in productivity, and will wrongly estimate the true output change if this is different from the change in inputs.

‘Volume indicators can relate to:

#### ‘Inputs

**C31** for example, the number of employees. This would simply assume that twice as large a public service would mean twice as much output, irrespective of how those additional personnel were deployed. The advantage of the method is the ease of implementation, and the ready availability of data. This method however ignores all changes in productivity due to, eg improved equipment (for example increased use of PCs) or more efficient procedures.

‘A possibility would be to complement input methods with adjustments for changes in productivity ...

‘The problem is that such adjustments are inevitably based on assumptions, which cannot be verified without genuine measurement of the output ...

‘Another problem is that there might be double counting of the productivity changes, if the quality changes of the inputs (eg the labour) were already taken into account ...

**‘Activity**

- C32** for example, number of operations in hospitals or number of patrols carried out by the police. Such data can often be found. Activity indicators reflect what the non-market units are actually doing with their inputs and are therefore closer to the output. However, suppose for example that new improved forms of medical treatments reduce the number of operations necessary. Taking the number of operations as an indicator would imply a decrease of output and productivity, which does not seem appropriate in this case. Using activity indicators often does not lead to reasonable productivity numbers. However, for some collective services, activity indicators may be the only indicators that can be found.

**‘Output**

- C33** the preferred approach. However, it is not always easy to define exactly what the unit of output is. For individual goods and services it is in principle possible to define the output, since an actual delivery of that output takes place from the producer to the consumer(s) ... For example, for education, the output is the amount of teaching consumed by a pupil. For hospital services, the output is the amount of care received by a patient. For cultural services, the output is the amount of theatre plays consumed. For collective services, however, there is no transaction between producer and consumer since these are provided simultaneously to the society as a whole. It becomes therefore very difficult to define the output. It is very difficult to say for example what the unit of output is of defence or police services.

**‘Outcomes**

- C34** for example, indicators of the level of education of the population, life expectancy, or level of crime. Such indicators might be influenced by factors that are unrelated to the activity, and therefore are generally not representative of the output. In some cases, however, outcome indicators can be used as indicators for the quality of the output ...’ (section 3.1.2.1)

Table C3 Summarised from section 3.1.2.3 of the Eurostat Price and Volume manual

Type of service	A/B/C methods
Individual services	<p>A methods – output indicator approach where the indicators satisfy the following criteria</p> <p>a) they should cover all the services produced by the producer that are provided to external users, and only those; activities that are in fact ancillary to the main output should not be counted;</p> <p>b) they should be weighted by the costs of each type of output in the base year;</p> <p>c) they should be defined as detailed as possible;</p> <p>d) they should be quality-adjusted.</p> <p>B methods – output indicator approach where the criteria are not fully satisfied eg. level of detail could be improved or does not take into account changes in quality.</p> <p>C methods – if input, activity or outcome is used (unless outcome can be interpreted as quality-adjusted output) or if coverage of output method is not representative.</p>
Collective services	<p>Broadly the same as for individual services but:</p> <p>B methods – input methods are B methods as are the use of volume indicators of activity. If input methods are used they should estimate the volume of each indicator separately, taking quality changes of inputs into account. For each category of inputs (IC, other taxes and subs on prod, comp of employ and CC). Applying productivity or quality adjustments to the sum of the volume of inputs is not recommended.</p> <p>C methods – the use of a single input volume indicator is not a B method.</p>

**C35** Specifically for Education, Eurostat proposes several methods of measuring output, all classified as either A, B or C method. The criteria for the different methods are set out in full in section 4.2 of the Eurostat Handbook. In summary they are as follows:

- ‘A method: complete or near-complete coverage, stratification by category, at least into pre-school, primary, lower secondary, upper secondary, higher education and other education. The A method for non-market education is to use “pupil hours” adjusted for quality as appropriate. The number of pupils can be used as a proxy for pupil hours provided that it can be shown that the amount of hours that pupils spend being taught is sufficiently stable. It is encouraged for some levels of education (tertiary education and distance-learning).
- ‘B method: complete or near-complete coverage stratification by category, at least into pre-school, primary, lower secondary, upper secondary, higher education and other education. The B method is to use “pupil hours” without an adjustment for quality.
- ‘C method: incomplete coverage or incomplete stratification by category. An input-base method is a C method, such as numbers of teacher hours.

### Measuring quality change

**C36** ‘When a constant price measure is not obtained by deflation with a price index but instead by extrapolation with a volume index, quality changes should also be accounted for. This however provides some special problems ... (section 2.4.3)

‘When volume indicators with a detailed breakdown of products are used, shifts between different products will be included in the volume components. Therefore, part of the quality change (that part due to compositional changes in an aggregate ...) can be captured by differentiating as many qualities of a product as possible. These different qualities are then in fact treated as different products.’ (section 2.4.3)

## Capital Measures and Definitions

### OECD Manual guidance on capital input measures

- C37** ‘For any given type of asset, there is a flow of productive services from the cumulative stock of past investments. This flow of productive services is called *capital services (sic)* of an asset type and is the appropriate measure of capital input for production and productivity analysis. Conceptually, capital services reflect a quantity, or physical concept, not to be confused with the value, or price concept of capital. To illustrate, take the example of an office building. Service flows of an office building are the protection against rain, the comfort and storage services that the building provides to personnel during a given period.

‘Because flows of the quantity of capital services are not usually directly observable, they have to be approximated by assuming that the service flows are in proportion to the stock of assets...’ (paragraph 5.2)

- C38** ‘The Australian Bureau of Statistics publishes two distinct and complementary capital measures. The Australian methodology stands out in that it ensures full consistency between the different measures:

- A measure of capital services, as part of ABS’ multifactor productivity series.
- An end-year net capital stock, as part of the Australian System of National Accounts.’ (Box 6, pp 60)

### SNA Guidance on the boundary between current and capital expenditure

- C39** ‘The distinction between maintenance and repairs and gross fixed capital formation is not clear-cut. The ordinary, regular maintenance and repair of a fixed asset used in production constitutes intermediate consumption. Ordinary maintenance and repair, including the replacement of defective parts, are typical ancillary activities but such services may also be provided by a separate establishment within the same enterprise or purchased from other enterprises.

‘The practical problem is to distinguish ordinary maintenance and repairs from major renovations, reconstructions or enlargements which go considerably beyond what is required simply to keep the fixed assets in good working order. Major renovations, reconstructions, or enlargements of existing fixed assets may enhance their efficiency or capacity or prolong their expected working lives, They must be treated as gross fixed capital formation if they add to the stock of fixed assets in existence.

- C40** ‘Ordinary maintenance and repairs are distinguished by two features; (paragraph 6.161)
- a) They are activities that owners or users of fixed assets are obliged to undertake periodically in order to be able to utilise such assets over their expected service lives ...

- b) Maintenance and repairs do not change the fixed asset or its performance, but simply maintain it in good working order or restore it to its previous condition in the event of a breakdown. Defective parts are replaced by new parts of the same kind without changing the basic nature of the fixed asset.

**C41** ‘On the other hand, major renovations or enlargements to fixed assets are distinguished by the following features:

- a) The decision to renovate, reconstruct or enlarge a fixed asset is a deliberate investment decision which may be taken at any time and is not dictated by the condition of the asset ...
- b) Major renovations or enlargements increase the performance or capacity of fixed assets or significantly extend their previously expected service lives ...

**C42** ‘Research and development are undertaken with the objective of improving efficiency or productivity or deriving other future benefits so that they are inherently investment – rather than consumption-type activities. However, other activities, such as staff training, market research or environmental protection, may have similar characteristics. In order to classify such activities as investment type it would be necessary to have clear criteria for delineating them from other activities, to be able to identify and classify the assets produced, to be able to value such assets in an economically meaningful way and to know the rate at which they depreciate over time. In practice it is difficult to meet all of these requirements. By convention, therefore, all the outputs produced by research and development, staff training, market research and similar activities are treated as being consumed as intermediate inputs even though some of them may bring future benefits.’

‘When an enterprise contracts an outside agency to undertake research and development, staff training, market research or similar activities on its behalf, the expenditures incurred by the enterprise are treated as purchases of services used for purposes of intermediate consumption.’ (paragraph 6.165)

### OECD Guidance on Productivity

**C43** Productivity is not defined (at present) within National Accounts guidance. The following extracts are from the OECD productivity manual.

**C44** ‘Emphasis is given to productivity measures of those industries that are characterised by a large share of market producers, leaving aside those activities where non-market producers dominate in many OECD countries. These activities pose specific problems of productivity measurement, due to the difficulty or impossibility of observing and/or defining market prices or output. Reference will be made when appropriate but an in-depth treatment of the output measurement in each of these industries would go beyond the scope of the present manual.’ (paragraph 1.2)

### Defining productivity

**C45** ‘Productivity is commonly defined as a ratio of a volume measure of output to a volume measure of input use ... there is neither a unique purpose for, nor a single measure of productivity’. The objectives of productivity measurement identified by the OECD manual are:

- to trace technical change;
- for identifying changes in efficiency;
- to identify real cost savings in production. (Where real cost savings are a result of the many factors that drive productivity growth including technical change and changes in efficiency);
- to help identify inefficiencies; and
- as a key element towards assessing standards of living.

Table C4 Overview of main productivity measures

Type of output measure	Type of input measure			
	Labour	Capital	Capital and Labour	Capital, labour and intermediate inputs (energy, materials, services)
Gross output	Labour productivity (based on gross output)	Capital productivity (based on gross output)	Capital-labour productivity MFP (based on gross output)	KLEMS multifactor productivity
Value added	Labour productivity (based on value added)	Capital productivity (based on value added)	Capital-labour productivity MFP (based on value added)	-
	Single factor productivity measures	Single factor productivity measures	Multifactor productivity (MFP) measures	Multifactor productivity (MFP) measures

Source: OECD Productivity Manual, Table 1

Table C5 Purposes, advantages and limitations of the most widely used productivity measures

Type of measure	Purposes	Advantages	Limitations
<b>Labour productivity based on gross output</b>	'Traces the labour requirements per unit of (physical) output.'	'Ease of measurement and readability. In particular, the gross-output measure requires only price indices on gross output, not on intermediate inputs as is the case for value-added based measures.'	'Labour productivity is a partial productivity measure and reflects the joint influence of a host of factors. It is easily misinterpreted as technical change or as the productivity of individuals in the labour force.'
<b>Labour productivity based on value added</b>	'Analysis of micro-macro links, such as the industry contribution to economy-wide labour productivity and economic growth... a direct link to a widely used measure of living standards, income per capita...'	'Ease of measurement and readability.'	As above, and 'Also, value-added measures based on a double deflation procedure with fixed-weight Laspeyres indices suffer from several theoretical and practical drawbacks'
<b>Capital-labour productivity MFP based on value added</b>	'Analysis of micro-macro links, such as the industry contribution to economy-wide MFP growth and living standards, analysis of structural change'	'Ease of aggregation across industries, simple conceptual link of industry-level MFP and aggregate MFP growth. Data directly available from national accounts'	'Not a good measure of technology shifts at the industry or firm level. When based on value added that has been double deflated with a fixed weight Laspeyres quantity index [sic], the measure suffers from the conceptual and empirical drawbacks of this concept'
<b>Capital productivity based on value added</b>	'Changes in capital productivity indicate the extent to which output growth can be achieved with lower welfare costs in the form of foregone consumption'	'Ease of readability'	'Capital productivity is a partial productivity measure and reflects the joint influence of a host of factors. There is sometimes confusion between rates of return on capital and capital productivity'.
<b>KLEMS multifactor productivity based on gross output</b>		'Conceptually, KLEMS-MFP is the most appropriate tool to measure technical change by industry as the role of intermediate inputs in production is fully acknowledged.'	'Significant data requirements, in particular timely availability of input-output tables that are consistent with national accounts'

Source: OECD Productivity Manual, 2001

### ESA Guidance on Satellite Accounts

- C46** ‘For some specific data needs the best solution is to draw up separate satellite accounts ...
- ‘Satellite accounts can serve such data needs by:
- a) showing more detail where necessary and leaving out superfluous detail;
  - b) enlarging the scope of the accounting framework by adding non-monetary information, e.g. on pollution and environmental assets;
  - c) changing some basic concepts, e.g. by enlarging the concept of capital formation by amount of the expenditure on research & development or the expenditure on education.
- C47** ‘An important feature of satellite accounts is that in principle all basic concepts and classifications of the standard framework are retained. Only when the specific purpose of satellite account definitely requires a modification, are changes in concepts introduced. In such instances, the satellite account should also contain a table showing the link between the major aggregates in the satellite account and those in the standard framework. In this way, the standard framework retains its role as a framework of reference and at the same time justice is done to more specific needs. (paragraph 1.20)
- C48** ‘The standard framework does not pay much attention to stocks and flows which are not readily observable in monetary terms... By their nature, the analysis of such stocks and flows is usually also well served by compiling statistics in non-monetary terms, e.g. (paragraph 1.21)
- ...b)education can be described in terms of type of education, number of pupils, the average number of years of education before obtaining a diploma, etc...
- ‘They [extended accounts] can also reclassify the final expenditure on regrettable necessities (e.g. defence) as intermediate consumption, i.e. as not contributing to welfare... In this way, one could try to construct a very rough and very imperfect indicator of changes in welfare. However, welfare has many dimensions, most of which are best not expressed in monetary terms. A better solution for measuring welfare if therefore to use, for each dimension, separate indicators and units of measurement. The indicators could be, for example, infant mortality, life expectancy, adult literacy and national income per capita. These indicators could be incorporated in a satellite account.’ (paragraphs 1.18-1.22)

# Appendix D: How OECD Countries Measure Government Output

**D1** In order to understand how government output is measured elsewhere, ONS conducted a survey of OECD countries in 2003. Those countries making most progress in direct volume output measurement were subsequently consulted about their methodologies. This appendix notes, to the best of our knowledge, the methodologies uncovered by the survey and consultations. The functions, by country, covered are:

**Health:** Australia, New Zealand, Italy, the Netherlands, Finland and Norway.

**Education:** Australia, Italy, the Netherlands, Finland, Germany, Norway, Sweden and the United States.

**Public Order and Safety:** Australia.

**Social Protection:** Finland.

**D2** Most countries still retain the old system of assuming government output as being equal to the value of the inputs. However, alongside the United Kingdom, some countries have been making progress in developing direct volume output measures for some functions of government. Australia, New Zealand, Italy and the Netherlands are the other front-runners, with direct volume measures covering between 20 and 50 per cent of government output. For all four countries, this is largely confined to Health and Education.

**D3** In line with the European Commission Decision of 17 December 2002, several European countries are also developing measures for incorporation in their national accounts over the next few years. These include Finland, Germany, Norway and Sweden, with the focus again being on Health and Education. The United States is also developing direct volume measures for Education output.

## Health:

**D4** Australia, New Zealand, Italy and the Netherlands already use direct volume measures for government Health output in the national accounts. Meanwhile, Finland and Norway are developing measures for implementation in 2006.

**D5** From the following summaries, we can see that most countries use B methods of output measurement, according to the Eurostat guidelines, with none yet successful in capturing quality change, and most using Diagnosis Related Groups (DRGs) to measure hospital treatments. On smaller components of the health system, B and C methods are also in use.

## Australia

**D6** Health outputs that are directly measured include: hospital and nursing home services; and services provided by GPs and medical specialists.

**D7** For hospital services, output is an index derived from numbers of treatments ('episodes') classified over 660 DRGs, weighted by costs derived from annual cost studies. For nursing home services, the measure is patient-days categorised by level of care offered. GP and specialist medical services are measured by numbers of attendance, weighted by fees charged.

**D8** Australian Bureau of Statistics (ABS) generally expects that, by weighting at a detailed level and using cost weights, its results will pick up quality change, as would be the case in the market sector. Estimates for Health capture quality and productivity changes through the compositional shifts between treatment types with differing costs, but not within a particular treatment type. In measuring treatments by GPs, the categories used for weighting reflect different levels of qualifications.

Remaining health services, such as dentistry, optometry, community health services, paramedical and ambulance services will continue to be estimated on an inputs basis.

### New Zealand

**D9** The Health indicator is based on case-mix adjusted inpatient discharges. Each treatment is classified to one of 667 DRGs, which are aggregated using base year average prices for each category. The case-mix will change each year.

### Italy

**D10** Hospital services and outpatient services are separated. Changes to the statistical measurement of hospital services in Italy have coincided with new funding methods. Hospitals are now funded on numbers of clinical cases by DRG; previously the basis was patient-days.

**D11** Hospital services are measured by patient numbers, sub-divided into treatment groups, and classified into about 20 broad clinical specialities. The volume measure by clinical group is adjusted by an 'average stay' indicator.

**D12** No explicit quality indicators are in use, although possibilities based on levels of technical facilities, and compliance with Ministry of Health qualitative standards were considered.

### The Netherlands

**D13** Hospital treatments are to be measured, generally, by patient discharges, by the Hospital Discharge Register (HDR), although in certain cases, notably repeat treatments of cancers, treatment courses will be separately recorded.

**D14** The volume measures are weighted by average number of hospital-days of treatment for in-patients, with 'day' treatments counted as one day. There are approximately 7,000 categories; i.e. 989 International Classification of Diseases (ICD) diagnoses by seven age groups.

**D15** The introduction of extra funding to reduce waiting lists in 2001 had led to a marked rise in measured output in that year. Data on outpatient treatment were weaker, with a particular problem being treatments given outside hospital by medical specialists whose fees, taken from the CTG tariffs, were not closely related to different treatment costs.

### Finland

- D16** Finland also plans to implement direct volume output measures for Health in 2006. Its system is divided into hospital services and health care services, run by 450 municipalities (a sample of 10–20 per cent of these is used for output measurement), while psychiatric care (with no output indicators as yet) is provided by central government.
- D17** Hospital services are being investigated using DRGs. For inpatients, a treated patient is regarded as the output, weighted according to cost share. For outpatients, where the treatment is unrelated to inpatient care, output is similarly classified and weighted. For the health centre, output is measured by the number of visits for each of 15 categories. Each healthcare centre has made an estimate of the unit cost for each category. These are used as weights.

### Norway

- D18** Norway plans to implement direct volume output measures for Health by 2006.
- D19** For hospital services, Statistics Norway hopes to measure DRGs with an adjustment for quality. The DRGs take into account new treatments and discard old and less significant treatments. Data collection on the number of treatments is good and detailed, as Statistics Norway is able to collect this direct from hospitals, by using registers.
- D20** However, due to the inability to attach costs, assigning appropriate weights to treatments remains problematic. Quality adjustment also remains an issue to be dealt with.

### Education:

- D21** Australia, Italy and the Netherlands already use direct volume output measures for government Education output in their national accounts, using the number of pupils, weighted by cost. The quality adjustment is measured in different ways. Meanwhile, Finland, Germany, Norway, Sweden and the United States are developing measures for implementation over the next few years.

### Australia

- D22** The volume measures for Education output are based mainly on annual student enrolments. Enrolments for each level of education are weighted together by the cost of providing those services. Student numbers for primary schools and secondary schools are converted to full-time equivalents. Module hours are available for vocational education and are used in preference to student numbers. For universities, full-time equivalent student numbers enrolled in each of eleven discipline groups are used as the output indicators for the tuition component. The university research component is estimated by weighting together data for the number of publications and student research completions.
- D23** The new output indicators do not capture any quality change over time in the education services provided. The ABS is investigating ways of incorporating quality adjustment factors into the education output measure. For example, class sizes and public examination results have been considered as indicators of change in the quality of the education service. Adjusting for class size has been widely rejected on the grounds that there does not appear to be an observable relationship – certainly not a linear one – between class size and the quality of services provided. Public examination results cannot be used unless the exam standards are maintained over time.

While results from standardised tests provide a better prospect, changes in scores over time could also reflect external factors – such as changes in the quality of home life – as well as changes in the quality of education services. Nevertheless, once a sufficient time series of such data becomes available and after taking into account external factors, it may be possible to at least discern the direction of change in quality.

### Italy

- D24** The Education measure has complete coverage. It is stratified by primary, secondary, further education and university. School education output is measured by the number of pupils, together with a quality adjustment based on class size.
- D25** Italian studies on a good level of education showed that class sizes were relevant. There were other relevant standards such as physical size of classrooms and resources available. Other possible measures of education quality or examination results were considered but not adopted because they were thought not to offer sufficiently reliable and stable measures over a long time period. The Italian Statistical Authority (ISTAT) considered using class size and teaching equipment, but finally decided to use only class size to measure quality.
- D26** ISTAT introduced a conversion factor into volume to take account of service quality. Overloading a public service will reduce quality from a certain point, so the measure of output should use a conversion function. Output is considered linear up to a certain level of class size, grows sub-linearly beyond that, and reaches a maximum level at the point at which overloading begins. The conversion factor is based on the median size of classes, stratified by region and type and is annually estimated. This transforms the actual number of pupils into a ‘standard number’ who benefit. For elementary level, the maximum number of pupils should be 25. For example, if 25 is the maximum for a class but there are 28 in the class, then there will be a lower number of standard children.
- D27** ISTAT recognises that this quality measure can be seen as subjective. It regards its method for measuring Education as a B method because of their concerns about their measure of quality and the use of number of pupils rather than pupil hours. This method is not available or appropriate for universities, where they use different courses and costs of degrees.

### The Netherlands

- D28** School education is measured by a weighted index of quantity measures across ten categories of education from primary up to university education, with weightings determined by costs in the previous period.
- D29** For primary, special and senior vocational (secondary) levels, quantity indicators are based on pupil numbers. For secondary education levels up to pre-university, the quantity measure is based on pupils or students ‘moving up’ each year – i.e. under the Dutch system, pupils or students who have reached the required minimum level of attainment in order to move up to the next academic year-level. These assessments are done at school level. Data on the number of pupils moving up, however, will no longer be available.

- D30** There are standardised tests at ages 11 (CITO) and at 16 in the Dutch system. Statistics Netherlands investigated the viability of results from these as quality indicators but were persuaded on further investigation – including representations by the examination authorities – that they would be unsuitable for a national accounts purpose. The proportion of pupils with particular results was constant year on year, suggesting either that there is deliberate marking to produce standard proportions, or that there is objectively no movement in education standards as measured by CITO. There was also concern that schools would teach to the exam tests, undermining their independence as progress measures.
- D31** For vocational courses, the quantity index is the number of pupils adjusted for the expected duration of study before certification (So, one student on a course nominally of three years, but which normally takes four years, is measured at 0.75). University education and ‘other’ education services are measured on an input basis.
- D32** Although the Netherlands has a UK-style schools inspection regime, with publicly available inspection reports, these do not offer categorical quality rankings akin to those of Ofsted (‘excellent’, ‘very good’ etc), and are therefore considered unusable for quality adjustment.

#### Finland

- D33** Finland plans to implement direct volume output measures for Education by 2006, with most of the information backdated to 1995.
- D34** Statistics Finland plans to use number of students, or lessons for school level education. International guidance prefers the use of student-hours, but this is not collected in Finland. For university, polytechnic and vocational education, it will use the number of credits obtained, as this is a better measure of education than number of students (students need credits to achieve modules).

There is no plan to include a measure of quality initially.

#### Germany

- D35** From 2005, Germany will be implementing direct volume output measures for Education, using learning hours as the indicator. However, a theoretical number, rather than the actual number of learning hours will be produced, as actual numbers need to be adjusted to take account of lessons which do not take place, absences due to illness, truancy, etc. and the necessary time series is not available.
- D36** The Federal Statistical Office (FSO) research has concluded that the only convincing approach to measuring pupil performance is to use standardised tests and use it as an indicator of the quality of the teaching. However, FSO has decided not to introduce a quality adjustment with the introduction of the direct volume output measure in 2005. The FSO research urged caution against using what may be arbitrary quality indicators – if there are any doubts, an adjustment for quality should be dispensed with.

**D37** FSO will continue further research beyond the scope of the initial project and focus on the following:

- a) measurement of quality;
- b) representative data on the number of lessons which are cancelled or missed;
- c) by-products (universities' research activities, teaching provided by interest groups and teaching hospitals, further training provided by companies in-house, etc.);
- d) information on pupil numbers to be based on data for the calendar year rather than the academic year;
- e) stratification and international comparability; and
- f) choice of index formula chain index.

### Norway

**D38** Norway plans to implement a direct volume output measure for Education shortly after Health is implemented in 2006.

**D39** A research study into human capital and the rate of return of the Norwegian higher education sector was undertaken in 2003. Statistics Norway followed the Jorgenson and Fraumeni (1989) approach which measured output in the education sector by the increase in the total discounted lifetime income that can be attributed to the education 'produced' in a given year. Statistics Norway modified this approach by, first, confining the estimates to the higher education sector (90 percent of Norwegians in the 15-19 age group enter higher education), and, second, excluding the value of non-market labour activities.

**D40** The new baseline estimates showed that value-added in the Norwegian higher education sector was eight times higher than the corresponding figure in the Norwegian national accounts. Replacing the standard national accounts figure by the baseline estimate would have increased the share of higher education in GDP from 1.0 to 7.3 per cent.

### Sweden

**D41** Sweden plans to implement direct volume output measures for Education by 2006.

**D42** Compulsory education in Sweden begins with primary education (ages 7-16), followed by secondary education (ages 16-19).

**D43** Statistics Sweden has begun a development project. For primary schools, it hopes to measure the number of pupil hours as the volume indicator of output and within this a further split of the number of pupils learning their native language and Swedish as a second language. In combination with the number of pupils taught, statistics Sweden is also considering three additional indicators:

- health services provided by schools;
- school transportation; and
- school meals.

**D44** Statistics Sweden hopes to address the issue of quality within primary schools by adjusting the volume indicators by either measuring:

- the number of pupils leaving grade 9 with no qualifications;
- the changes in the achievement levels; or
- the number of pupils eligible for secondary education and changes in the achievement levels.

**D45** For secondary schools, Statistics Sweden hopes to measure the same volume indicators as for primary schools. However, data on pupil hours are not available, so the number of pupils for different subjects would be used.

**D46** Possible quality measures that Statistics Sweden is analysing include:

- changes in average grades;
- changes in average grade and pupil leaving with no qualifications;
- number of pupils graduating; and
- number of pupils leaving with other qualifications.

**D47** Preliminary work has begun on tertiary education. Statistics Sweden has been unable to obtain cost data and indicators for research and research education, i.e. active doctoral students or FTE doctoral students.

For a measure of quality, Statistics Sweden has investigated the following areas:

- exam results;
- quality checks made by the Swedish National Agency of Higher Education;
- performance points; or
- some other outcome related adjustment.

### The United States

**D48** The United States Bureau of Economic Analysis (BEA) is examining a number of possible ways of measuring output of the government sector. BEA has begun a research project to analyse the effects of using the number of pupils enrolled as the base index for primary and secondary education and to then consider various quality adjustments to this index. These quality adjustments include teacher quality, pupil-teacher ratio and high-school dropouts.

**D49** The preliminary results show that quality adjusting a volume indicator can have a significant effect on the estimated output. The difference between the annual growth rate for the quality-unadjusted measure and the preferred quality-adjusted measure (using teacher quality and pupil-teacher ratio) is 0.15 per cent for 1980-2001, 0.53 per cent for 1980-1990 and -0.20 per cent for 1990-2001, respectively.

### Public Order and Safety:

- D50** There is little to report on the progress of other countries in developing direct volume output measures for Public Order and Safety. Australia has recently abandoned attempts to measure public order functions in this way.

#### Australia

- D51** ABS has done some work in recent years on developing output measures for police services, justice services and corrective services. Its initial recommendations were published in a report in November 2000.
- D52** An output measure for Police was constructed, drawing primarily on ABS data on recorded crime. However, the measure was not considered suitable for implementation due to the number of assumptions required for the index to be representative of the sector. The report recommended that an input-based measure be retained for Police, but the report did suggest the kind of data and methodology that would be required in future for a credible output-based measure of Police.
- D53** Data limitations also thwarted the construction of a credible output measure for justice services. There were no data for parts of the sector, particularly tribunals and specialist courts and where data did exist, they did not include sufficient detail to allow differentiation between different case finalisation methods or case types within a court.

### Social Protection:

- D54** There has been little progress internationally towards direct measurement methods for Personal Social Services and Social Security. In certain countries with insurance-based regimes for health and social care, social care with accommodation is measured by deflating expenditures by standard insurance rates for accommodation. Social Security Administration is typically included under the heading of Public Administration and measured on an inputs basis. Only Finland appears to be currently developing direct volume output measures for these functions.

#### Finland

- D55** Finland plans to implement direct volume output measures for Personal Social Services and Social Security in 2006.
- D56** Personal Social Services, as for the United Kingdom, is split into several groups, including elderly care, children care, and welfare services. If the service is an outpatient one, the total number of visits is the indicator. If the service includes accommodation, then the number of occupant days is used. Quality change has not yet been considered.
- D57** Social Security, again like the United Kingdom, will be measured using the number of people receiving benefits by group (including pensions, disability and unemployment allowances), weighted by the cost shares of each.

# Glossary

---

**Blue Book:** the common name for the annual UK National Accounts.

**Compensation of Employees:** wages and salaries, national insurance contributions, pension contributions and redundancy payments paid by government.

**Chain-linking:** technique now adopted by ONS to construct volume data. Under previous methods volume data was constructed by measuring activity in the price of a specific year (the price reference base) that was updated about every five years. Under chain-linking the price reference base is updated annually. The technique better reflects changing patterns of output and expenditure.

**Devolved administrations:** the Scottish Parliament for Scotland, the Welsh Assembly for Wales and for Northern Ireland.

**Deflation:** the technique used to change figures from nominal terms (current prices) into real terms (constant prices or volume terms), expressing the production (or consumption) of goods and services in the prices of a common year.

**European System of Accounts (ESA 1995):** EU framework for National Accounts measurement to which the UK National Accounts adhere, currently ESA 1995 (see also SNA).

**Excellence in Cities:** combination of initiatives introduced by the government in March 1999 to tackle particular problems facing children in cities, eg to tackle disaffection, social exclusion, truancy and indiscipline and to raise aspirations and achievements and to improve parents' confidence in cities.

**Final Consumption Expenditure:** total cost of inputs used in the production of the final output.

**General Government (GG):** consists of the following group of resident institutional units:

- All units of central, state or local government;
- All social security funds at each level of government;
- All non-market non-profit institutions that are controlled and mainly financed by government units.

The sector does not include public corporations, even when all the equity of such corporations is owned by government units. It also does not include quasi-corporations that are owned and controlled by government units. However, unincorporated enterprises owned by government units that are not quasi-corporations remain integral parts.

**General Practice Research Database (GPRD):** the world's largest computerised database of anonymised clinical records which can be used to optimise decision-making across the health research spectrum including clinical epidemiology, drug safety, disease management and drug utilisation. It will be licensed to academics, regulators, pharmaceutical organisations and research service providers.

**Great Britain:** the countries of the United Kingdom without Northern Ireland.

**Gross Domestic Product (GDP):** headline measure of economic activity.

**Gross Fixed Capital Formation (GFCF):** investment in capital equipment – (by households, government or business).

**Gross Operating Surplus (GOS):** Gross Trading Profits of companies (and surpluses of public sector/government bodies) plus rental income on buildings.

**Gross Value Added (GVA):** Gross Domestic Product excluding taxes (less subsidies) on products; the headline measure of regional economic activity.

**Intermediate Consumption Expenditure:** cost of goods and services used in the production of the final output.

**Key Stages:** The National Curriculum is divided into four key stages according to pupils' ages. There are formal curricula for what must be taught in the different Key Stages.

- Key Stage One – Primary School (5-7 years old). Years 1-2.
- Key Stage Two – Primary and Secondary School (7-11 years). Years 3-6.
- Key Stage Three – Secondary School (11-14 years). Years 7-9.
- Key Stage Four (GCSE) – Secondary School (14-16 years). Years 10-11.

**Labour Force Survey (LFS):** a monthly survey of households which is the primary source of labour market information, including employment, unemployment and inactivity.

**New Earnings Survey (NES):** annual survey of the earnings of employees.

**Non-Market Output:** consists of goods and individual or collective services produced by non-profit institutions serving households (NPISHs) or government that are supplied free, or at prices that are not economically significant, to other institutional units or the community as a whole.

**Output = Input:** the convention by which the volume of government output is deemed to be equal to the volume of the inputs.

**Producer Price Indices (PPI):** monthly estimates of the prices of goods bought and sold by UK manufacturers.

**Public Expenditure Statistical Analyses (PESA):** an annual publication by the Treasury giving a detailed breakdown of public expenditure outturns and plans, including an estimate of spending for the regions and devolved administrations.

**Public Service Agreement:** an agreement between a government department and the Treasury, as part of the Spending Review, including objectives and targets.

**Productivity:** commonly defined as a ratio of a volume measure of output to a volume measure of input.

**The Royal College of General Practitioners (RCGP):** is the academic organisation in the United Kingdom for general practitioners. Its aim is to encourage and maintain the highest standards of general medical practice and act as the 'voice' of general practitioners on education, training and standards issues.

**System of National Accounts (SNA 1993):** international framework for measurement of National Accounts from which the ESA is derived (see ESA).

**Total Factor Productivity (TFP):** measure of the efficiency with which inputs are used to produce output; estimated as a residual in the growth accounting framework.

**Workforce Jobs (WFJ):** an estimate of employment, including breakdown by industry, compiled mainly from business surveys.

# List of Abbreviations

---

ABS	Australian Bureau of Statistics
BLS	Bureau of Labor Statistics
CABG	Coronary Artery Bypass Graft
CJS	Criminal Justice System
COFOG	Classification of the Functions of Government
CPI	Consumer Prices Index
CSO	Central Statistical Office
DCA	Department for Constitutional Affairs
DCMS	Department for Culture, Media and Sport
DEFRA	Department for Environment, Food and Rural Affairs
DfES	Department for Education and Skills
DH	Department of Health
DRG	Diagnosis Related Group
DTI	Department of Trade and Industry
DWP	Department for Work and Pensions
ESA	European System of Accounts
EU	European Union
FRA	Fire and Rescue Authority
FRSIT	Fire and Rescue Service Improvement Team
FTE	Full Time Equivalent
GCSE	General Certificate in Secondary Education
GDP	Gross Domestic Product
GGFC	General Government Final Consumption
GHS	General Household Survey
GOS	Gross Operating Surplus
GP	General Practitioner
GPRD	General Practice Research Database
GVA	Gross Value Added
HES	Hospital Episode Statistic
HMCI	Her Majesty's Chief Inspector of Schools in England
ICT	Information and Communication Technology
IR	Inland Revenue
IT	Information Technology
KS	Key Stage
LEA	Local Education Authority
LFS	Labour Force Survey
MFP	Multi Factor Productivity
MoD	Ministry of Defence
MPC	Monetary Policy Committee
NAO	National Audit Office

NI	National Insurance
NIESR	National Institute of Economic and Social Research
NES	New Earnings Survey
NHS	National Health Services
NPISH	Non Profit Institutions serving Households
ODPM	Office of the Deputy Prime Minister
OECD	Organisation for Economic Co-operation and Development
Ofsted	Office for Standards in Education
ONS	Office for National Statistics
PAF	Performance Assessment Framework
PESA	Public Expenditure Statistical Analyses
POS	Public Order and Safety
PPAF	Policing Performance Assessment Framework
PPI	Producer Price Index
PSA	Public Service Agreement
PSR	Pre-Sentence Report
PSS	Personal Social Services
PTCA	Percutaneous Transluminal Coronary Angioplasty
QALY	Quality Adjusted Life Year
RAP	Referrals, Assessments and Packages of Care
RCGP	Royal College of General Practitioners
RPI	Retail Prices Index
RPIX	Retail Prices Index excluding mortgage interest payments
SDS	Single Data System
SERPS	State Earnings Related Pension Scheme
SNA	System of National Accounts
TFP	Total Factor Productivity
VAT	Value Added Tax
WFJ	Workforce Jobs

## References

---

- Ahlroth, S., Bjorklund, A., Forslund, A., 1997, 'The output of the Swedish Education sector' *Review of Income and Wealth*, Vol. 43, pp. 89–104
- Ahmad, N, 2003, *Measuring Investment in Software*: OECD, STI Working Paper 2003/6
- Allsopp, C, 2004, Review of Statistics for Economic Policymaking, HM Treasury
- Australian Bureau of Statistics, 2002, 'New volume estimates for health and education services' Australia Now, *Year Book Australia 2002*
- Ashaye, T, Nov 2001, Recent developments in the measurement of general government output, *Economic Trends*.
- Beales, R E, 1967, *Estimates of expenditure, income and product at constant prices*, IARIW conference, Maynooth
- Blatchford, P, Goldstein, H, Martin, C. and Browne, 2002 in 'A Study of Class Size Effects in English School Reception Year Classes' *British Education Research Journal*, Vol 28
- Central Statistical Office, 1956, *National Income Statistics: Sources and Methods*, HMSO, London.
- Caplan, D, Feb 1998, The measurements of real public sector output in the National Accounts, *Economic Trends*.
- Caplan, D, Oct 1998, 'Measuring the output of non-market services', *Economic Trends*.
- Commission Decision 98/715 on measuring prices and volumes in ESA, *Eurostat*.
- CSO, 1985, *United Kingdom National Accounts: Sources and Methods*, HMSO, London
- Dawson et al, 2004, 'Developing new approaches to measuring NHS outputs and productivity', *CHE Technical Papers Series No. 31*.
- Department for Education and Skills: 'Every Child Matters', Green Paper, 2003.
- Department of Health, Social Services Performance Assessment Framework Indicators, 2002-03.  
<http://www.publications.doh.gov.uk/paf/#pub>
- Department of Work and Pensions, Departmental Report 2004.  
<http://www.dwp.gov.uk/publications/dwp/2004/dr04/pdf/dwpdr04.pdf>
- Drèze, J P and Stern, N H, 1987, 'The Theory of Cost-Benefit Analysis' in A J Auerbach and M S Feldstein, editors, *Handbook of Public Economics*, vol 2, North-Holland, Amsterdam.
- Eisner, R, 1988, 'Extended Accounts for National Income and Product', *Journal of Economic Literature*, vol 26: pp 1611–1684.
- Ervil A, et al, 2003, 'A theory-based measure of the output of the education sector' Statistics Norway
- European Systems of Accounts, 1995, *Eurostat*

- Handbook of Price and Volume measures of National Account, 2001, *Eurostat*
- Fisk, D and Forte, D, 1997, 'The Federal Productivity Measurement Program: Final Results', *Monthly Labor Review*, May: pp 19–28.
- Fraumeni, B, *et al*, 2004, 'Real output measures for the education function of government: A first look at primary and secondary education' Bureau of Economic Analysis
- Goldstein, H. *et al*. 2000, 'The use of value added information in judging school performance', University of London
- Goldstein, H., 2003, 'Measuring the performance of the educational system', University of London
- Gravelle and Smith reference is 'Discounting for Health Effects in Cost-Benefit and Cost-Effectiveness Analysis' *Health Economics*, volume 10, pp 587–599.
- Hicks, J, R 1940, The Valuation of Social Income, *Economica*, vol 7 pp 105–124
- Hulten, 1984, 'Productivity Change in State and Local Governments', *Review of Economics and Statistics*, vol 66: pp 256–266.
- Inflation Report, Feb 2004, *Bank of England*
- Inflation Report, May 2004, *Bank of England*
- Jorgenson, D.W., Fraumeni, B, 1989, 'The output of the Education sector' In Zvi Griliches (ed.) *Output Measurement in the Service Sector*. Chicago: National Bureau of Economic Research, vol 56, University of Chicago Press
- Jorgenson, D, 1963, Capital Theory and Investment Behaviour, *American Economic Review*, vol 53 pp 247–259, and 1995, *Productivity*, vol 1 and 2, MIT Press, Cambridge
- Le Grand, J, 1982, The Strategy of Equality.
- Lévytt, M S and Joyce, M A S, 1987, *The Growth and Efficiency of Public Spending*, Cambridge University Press, Cambridge.
- Mai, N, 2004, *Measuring health output in the UK: a diagnosis based approach*, ONS
- Maurice, R, editor, 1968, *National Income Statistics: Sources and Methods*, HMSO, London.
- McLean I., *et al*, 2003 *Identifying the flow of domestic and European expenditure into the English regions*, ODPM.
- NAO, 2003, *Making a difference: Performance of maintained secondary schools in England*, National Audit Office.
- OECD, 2001, *Measuring Productivity Manual*
- OECD, 2002, Education at a Glance, *OECD Indicators 2002*
- OFSTED, 1995, 'Class Size and the Quality of Education'
- O'Mahony, M., Stevens, P., 2004, *Outcome based measure for Education*, NIESR Conference
- Pritchard, A, May 2001, Measuring productivity in the provision of public services, *Economic Trends*.
- Pritchard, A, May 2002, Measuring productivity in the provision of public services, *Economic Trends*.

- Pritchard, A, July 2003, Understanding Government Output and Productivity, *Economic Trends*.
- Samuelson, P A, 1950, 'Evaluation of Real National Income', *Oxford Economic Papers*, vol 2: pp 1–29.
- Samuelson, P A, 1954, 'The Pure Theory of Public Expenditure', *Review of Economics and Statistics*, vol 36: pp 387–389.
- Schreyer, P, 2003, *Capital Stocks, Capital Services and Multi-factor Productivity Measures*, OECD
- Statistics Commission, *Revisions to Economics Statistics* (2004), Report no. 17, vol 1.
- Systems of National Accounts, *United Nations et al*, 1993
- UK National Accounts - *Blue Book* 1998
- UK National Accounts - *Blue Book* 2002
- UK National Accounts - *Blue Book* 2003
- Vignoles, A. et al. 2000, *The relationship between resource allocation and pupil attainment*, London School of Economics.

ISBN 0-11-703447-9



9 780117 034471