



**Economic and Social
Council**

Distr.
GENERAL

CES/2004/22
31 March 2004

Original: ENGLISH

STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

Fifty-second plenary session
(Paris, 8-10 June 2004)

**THE ATKINSON REVIEW OF UK GOVERNMENT OUTPUT MEASUREMENT – A
PROGRESS REPORT**

Invited paper submitted by Office for National Statistics, United Kingdom¹

INTRODUCTION

1. Len Cook, the United Kingdom National Statistician, recently commissioned an external review of how the Office for National Statistics measures government output and productivity². Government output accounts for 20 per cent of GDP, so measuring it accurately is not only important in itself but also important for the measurement of the economy as a whole. The Review is also concerned with the implied deflators and implied productivity estimates which are a by-product of the published output measures.
2. The Review is being undertaken by Sir Tony Atkinson of Oxford University, an economist who is internationally recognised for his work on public finance issues. He is supported by a team led by Joe Grice, on secondment from the Treasury, who is also the Deputy Head of the Government Economic Service.
3. A key objective of the Review is to recommend methods and approaches, which can be used to improve the measurement of UK government output. The terms of reference also cover the estimation of government productivity. The plans are to publish an interim report in July 2004 and the final report in January 2005.

MEASURING THE VOLUME OF GOVERNMENT SERVICES AS A COMPONENT

¹ Paper prepared by Alwyn Pritchard.

² <http://www.statistics.gov.uk/pdfdir/nsr1203.pdf>

OF GROSS DOMESTIC PRODUCT: RECENT HISTORY OF CONCEPTS AND METHODS

4. Measuring government services carries additional challenges as compared to measuring market output. Where goods and services are produced, sold and bought in a market situation, prices are charged. A business's production, for example, can be expressed in volume terms by stripping out the impact of changes in the prices obtained for it; that is not difficult as the prices are usually observable. Government services are mostly provided either free of charge or at a nominal charge: without observed prices, we are unable to measure volume in the way we usually do in the market sector.

5. How then can we measure the production of government services in volume terms? The earliest guidance, in the System of National Accounts (SNA), 1968 version, took the view that estimates of volume changes in non-market output should reflect changes in the inputs used up, on the grounds that it was "questionable whether reliable adjustments can be made for other factors contributing to labour productivity"³.

6. The revision of the SNA in 1993⁴ took a different line and concluded that government output should, where possible, be measured in terms of the outputs produced. It went on: "The object is to measure the quantities of services actually delivered to households"⁵. Later, it adds: "There is no mystique about non-market health or education services which make changes in their volume more difficult to measure than volume changes for other types of output, such as financial or business services"⁶. Though admitting that collective services posed difficulties, the advice for all other non-market services was to add together the numbers of units of outputs produced and to weight them using their relative costs of production.

UK IMPLEMENTATION OF SNA 1993 / ESA 1995

7. The 1993 revision of the SNA was welcomed by the ONS: it seemed clear to us that measuring outputs directly was methodologically superior to measuring inputs and using that as a proxy for outputs. Also, key users of our statistics took the same view which was particularly appropriate at a time when increasing interest was being taken in recording the performance of government. ONS therefore put in place a plan to measure government outputs and implement other new features of the new SNA / ESA.

8. A measure of government output was required for inclusion in two separate locations in the national accounts:

- as a component of gross domestic product on the production measure (and estimated as the value added by government to its inputs); and
- as a component of the gross domestic product on the final expenditure measure (and estimated as general government final consumption in volume terms).

The second of these corresponds to the full sales value of the production and is therefore larger than the value added. Government therefore has a lower weight in GDP (Production) than in GDP (Final expenditure). This development work made use of the experience gained in Task

³ System of National Accounts, 1968, §8.51 (available at <http://unstats.un.org/unsd/sna1993/toctop.asp>)

⁴ System of National Accounts 1993, § 16.133

⁵ System of National Accounts 1993, §16.135

⁶ System of National Accounts 1993, §16.138

Forces convened by Eurostat and from the guidelines being developed for producing output measures in Eurostat's Handbook on Measuring Prices and Volumes⁷.

9. In 1998, ONS began to produce output volume series for a few government functions – education, healthcare and social security. These were used as components of the published GDP estimate. Details of the indicators used have been published⁸ and the methods used to produce them are in line with the conclusions of the Eurostat Task Forces and with current methods used in many other countries. They were:

- in education, numbers of pupils being taught, weighted according to the overall national costs of different levels of course (nursery, primary and secondary);
- in health, a count of treatments broken down into a few high level categories, each weighted according to the relative costs of each treatment type;
- in social security, numbers of administrative processes carried out, weighted according to the cost of each type of process.

10. The education output measure was further adjusted to reflect evidence that the quality of education had been improving; the adjustment was based on the improvement in examination performance over a number of years and hence assumed that the improvement would continue annually at a constant rate. No separate quality adjustment was made for either health or social security. The outputs of the remaining government functions continued to be measured in terms of the inputs used.

11. Table 1 shows some of the estimates of government outputs which have been published. These volume series are based on the final expenditure measure of GDP. Also shown in this table are series showing volume of inputs (“real expenditure”). (To avoid confusion, the figures shown are those currently in use rather than the ones published at that time.)

⁷ Handbook on Price and Volume Measures in National Accounts, Eurostat, 2001, available at:

<http://europa.eu.int/comm/eurostat/Public/datashop/print-catalogue/EN?catalogue=Eurostat&product=KS-41-01-543--N-EN>

⁸ “Measuring the output of non-market services”, David Caplan, *Economic Trends*, October 1998

Table 1: Consumption expenditures incurred by general government, United Kingdom

Expressed as volume of inputs and volume of outputs, index numbers, 1995=100⁹

	Healthcare		Education		Social Protection	
	<i>inputs</i>	<i>outputs</i>	<i>inputs</i>	<i>outputs</i>	<i>inputs</i>	<i>outputs</i>
1995	100	100	100	100	100	100
1996	103	103	98	101	103	101
1997	104	105	98	102	103	103
1998	108	108	99	103	103	103
1999	115	110	101	104	108	104
2000	119	111	104	104	112	104
2001	126	113	112	105	111	105
2002	131	116	118	105	130	105

As noted before, these figures will necessarily be of importance – both in the context of assessing the output of key public services and as a contribution to GDP growth.

THE PROCESS OF THE ATKINSON REVIEW

12. Sir Tony Atkinson has overall responsibility for conducting the Review. He has a small team from across government and the Bank of England assisting him. The Review will report to the National Statistician who will decide how to respond to its recommendations. The process is independent of Government though, like other users and commentators, government departments are taking a close interest in it.

13. The aim of the Review is “to advance methodologies for the measurement of government output, productivity and associated price indices”. While, as noted above, there will be an interim and final report, there is no need to wait for the end of the Review before making possible improvements. On the contrary, where progress can be made quickly, the Review team will work closely with the National Statistician and ONS to ensure its implementation. Other improvements will naturally take more time to introduce.

14. Although this paper was written shortly after the start of the Review, a number of specific issues have already emerged as worthy of further attention. Some of these are considered in the next few paragraphs.

DEFINITION OF UNITS OF OUTPUT AND COVERAGE ISSUES

15. There remain significant issues as to how best to measure the outputs of public services and other parts of government. The overriding principle is that we are looking for measures of the value added by government spending, just as the other components of GDP represent value added in the economy. In turn, and in accordance with the international guidelines, it is not

⁹ Output indices derived from *UK National Accounts, The Blue Book 2003*, Table 6.5; input indices based on updating of Table 3 of “Understanding government output and productivity”, Alwyn Pritchard, *Economic Trends*, May 2003; available at: <http://www.nationalstatistics.gov.uk/cgi/article.asp?id=400>

therefore enough to have indicators of activity alone. So the number of patients treated, for example, will not be an ideal measure. We need to assess what is the value that those patients receive from being treated. Similarly, the number of pupils taught can only at best be a starting point. Again, we need to determine the value that pupils and society obtain from those lessons being taught. Outcomes are therefore an important dimension to consider. Clearly, we must be careful not to attribute to government spending social outcomes, which are actually the result of factors, which have nothing to do with the relevant public spending. People can become healthier and live longer for reasons which have little or nothing to do with health service expenditure or activities. But where changes in outcomes do result directly from government spending, they are undoubtedly relevant.

16. This focus on value in measuring output means that the quality dimension is an integral part of the issue. Measuring changes in the quality of public services is certainly not straightforward. But a number of sources of evidence are available to begin to form a balanced assessment. For example, there is a cottage industry assessing the clinical results of various forms of healthcare. These suggest ongoing improvements over time. Furthermore, there is substantial evidence available on patients' perceptions of the treatments they receive. That also seems a relevant source of information on the value they perceive themselves as receiving, which has a bearing on what they would be willing to pay for that treatment if need be. Furthermore, in the UK at least, there is growing material available on independent inspection and rating of various healthcare units. Changes in these ratings over time offer a further clue to movements in quality, to be taken into account with other information.

17. Focusing on value is also consistent with the view in the international guidance that the stress should be on the viewpoint of the user of the service rather than on the producer. This has some important implications. For example, healthcare suppliers have a tendency to see a new form of treatment – the result of technical progress - as a different output to the one it replaced, perhaps because the former is carried out in a different type of healthcare setting with a separate accounting system. In national accounts, we need to define units of output in a way that is recognisable to consumers as well as to suppliers. This could mean treating different processes as producing the same output so long as they bring the same result to the patient. Adopting this method would give different results to the supplier-based approach.

OUTPUTS OF COLLECTIVE SERVICES

18. There has been less agreement on – and perhaps less discussion of – what is the output of collective services such as police, fire services and defence. Although the SNA does not yet recommend measuring the output of collective services directly, nevertheless these are challenges that need to be faced. Citizens and taxpayers whose money is spent on such services can expect to have some indication of their output and productivity trends. If national accountants shy away from this, others may fill the void, perhaps less rigorously.

19. The focus on the consumer rather than the producer has important implications here. First, it argues in favour of an “option-based” approach rather than a cost weighted activity. The fire service is a case in point. It is easy to measure the activities of fire fighters using a cost-weighted activity index and to conclude that the more fires they attend, the more these activities increase GDP. Let us consider how consumers see it. Actually, most people's homes did not burn down last night, just as they do not every night. But citizens will widely have

enjoyed the benefits of having the fire service in place; their benefit comes from having the option of calling out the fire brigade, should – unusually and unfortunately - their house catch fire.

20. Secondly, the value of the option has to take account of the fact that what is being protected by fire fighters varies between households and will also vary over time independently of the inputs of the fire service. If I bought expensive new furniture this year, the protection I am getting from the existence of an efficient neighbourhood fire service has increased. The contribution of insurance companies to GDP would reflect this situation.

THE “TRIANGULATION PRINCIPLE”

21. National accounts are at their most robust when estimates are supported by a wide range of interlinking evidence. Where direct and perfect measurement is not feasible, we often rely upon indicators, at least in the short term. Confidence in the story they are telling is naturally increased if other available evidence corroborates the story. If it does not - and points in a different direction – then that seems like a warning bell that needs to be followed up.

22. The desirability of corroborational evidence is at least as strong in the context of measuring public sector output as in other sectors. In this context, the Triangulation Principle is one to be borne in mind. As a matter of arithmetic, inputs, outputs and productivity form a triangle. Productivity is the ratio of outputs to inputs but that identity does not imply that productivity should always be regarded as the arithmetic residual. It may well be that information is separately available about the productivity of a service independently of what the output indicators show. That evidence needs to be taken into account. How much relative weight is placed upon it will depend upon how much confidence we have in the output indicators themselves. If the output indicators are weak or felt to be unreliable more weight would naturally be placed on other sources of evidence about productivity in deciding the estimate for both productivity and output.

THE BOUNDARY BETWEEN CAPITAL AND CURRENT EXPENDITURE

23. The Review is also likely to examine the boundary between government expenditure which directly supports current production and that which does not. There are two dimensions to this.

24. First, there are longstanding issues as to where precisely the boundary should be drawn. For example, the SNA 1993 touches on the role of research and development: the current guidance is to treat it as intermediate consumption¹⁰. As such, research and development is an extra cost in the year it took place although there is an expectation that it will result in an increase in future output. Even though spending on research and development creates no net output for government in the year it takes place, there is nevertheless a case for distinguishing it in order to gain a better understanding of government output as a production measure, particularly when the expenditure on it is uneven. This is one of the topics being reviewed ahead of the next revision of the *System of National Accounts*. But the same argument applies to other types of spending, too. For example, the same argument applies to the substantial amounts government spends on training staff and on some forms of maintenance, both of which may have

¹⁰ System of National Accounts 1993, §6.163

similar characteristics to research and development.

25. Discussions to change the SNA necessarily and properly take time. Even for spending which now and in the future is properly classified as current, transparency and provision of data valuable to users would suggest distinguishing between current spending which directly supports current activity and current spending which is intended to build up a capability for the future. Current spending on clinical governance to improve the treatment of patients now and for the future, for example, is different in kind from current spending directly on the treatment of current patients. The presumption is that transparent statistics ought to show this, so that users can draw informed conclusions.

TIMELINESS AND FREQUENCY OF RESULTS

26. ONS has devoted much effort to defining concepts and finding the means to measure them sufficiently accurately. Doing this in a way which fits in with the timetable for producing GDP (Production) and GDP (Expenditure) is a challenging assignment: we need quarterly results and they must be available two months after the end of each quarter. What our development work has yielded to date is indicators which are data intensive to construct and which cannot be constructed quickly enough to meet the timetable; in many cases, there is a lag of two years before these results are available. There is a need now to move to a two level approach. At one level, we need a benchmark measure estimated when complete data are available, if necessary, computed only annually if that is the only way to ensure reliable measurement. At the other level, we need to identify indicators which become available quickly – less than two months after the end of a quarter – and which track the benchmark acceptably well. A processing system should be created which computes results from these two levels. The data collection should be accompanied by a constant monitoring of the production operations of government so as to pick up evidence to corroborate the statistical output measure. This is necessary in order to ensure that our figures are believed by those whose output is being measured and by other users.

INTERNATIONAL BEST PRACTICE

27. In order to ensure that the UK works in line with the best international practice, the Atkinson Review team is collecting information about how other countries are measuring government output. Visits are being made to some of these countries to study their methods, the data sources used to compile their government output estimates and to hear the views of their statistics users. The Review team is also closely in touch with the OECD and Eurostat, for the same reasons. These experiences will be taken into account in the Review's recommendations.

RELATIONSHIPS WITH THE REST OF GOVERNMENT

28. As noted earlier, government departments, like other users, are showing a high degree of interest in the Review's work. This is natural and reassuring at a time when an output based approach to public services is increasingly becoming the norm, and when in the UK at least, public services are very high on the domestic political agenda.

29. But public services such as health or education are complex. No single measure of their output, however well-based and considered, can do justice to this complexity. Official statisticians have a role to play, therefore, not only in producing reliable estimates of public output for the national accounts but also in producing and relating those estimates to a wider range of authoritative measures that describe the performance of those services, or at least which provide users and commentators with the means to make their own assessment.

WHAT WE HOPE TO ACHIEVE

30. The spotlight is currently on the measurement of government output in the UK. The Atkinson review gives an ideal opportunity to build up these statistics so as to achieve a higher level of credibility and understanding. The spotlight also means, however, that there is an onus to deliver. Some improvements can be made quite quickly: others will take longer. But the work is underway and the Atkinson Team welcomes ideas and collaboration from anyone who wishes to contribute.