How to Measure Government Productivity: A Review Article on 'Measurement of Government Output and Productivity for the National Accounts' (The Atkinson Report)

Aled ab Iorwerth[1]
Finance Canada

ABSTRACT

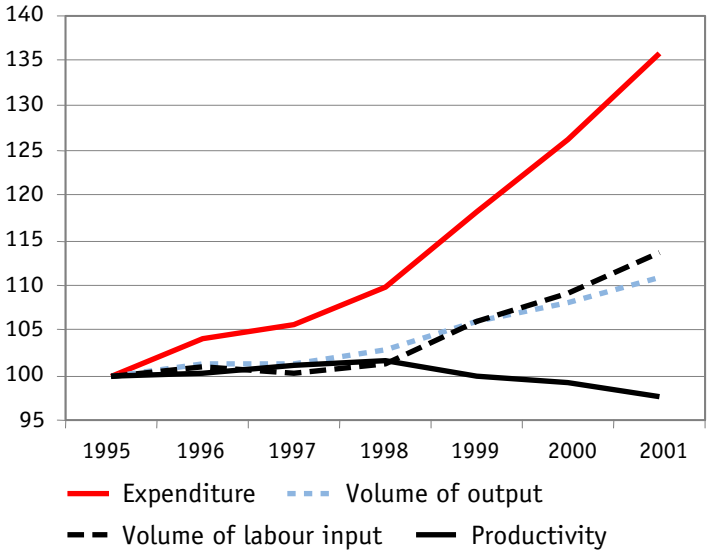
The Atkinson Report examined measurement of public services' output and productivity. Traditionally, only inputs to the government sector were reflected in the National Accounts. This treatment imposed zero productivity growth for the government sector. From the late 1990s, the United Kingdom — in accordance with the recommendations of international standards — had started to introduce direct measures of government output. Rather than count the number of teachers, a measure of the education sector's output would include pupil attendance in schools, for example. Introducing these measures of government output had cast light on what the government sector produces and its productivity in doing so. Sir Tony Atkinson reviewed these methods and recommended improvements. The objectives of introducing direct measures of government output and the principles he advocated in undertaking this task are summarized in this review. As well as ensuring an important part of the economy is reflected in the National Accounts, introducing direct measures of government output into the National Accounts would allow policymakers to make more informed judgments and give greater information on the overall performance of government services.

PROFESSOR ANTHONY ATKINSON WAS asked in December 2003 by the Office for National Statistics (ONS) in the United Kingdom to prepare a report on the use of direct measures of government output in the National Accounts. The value added of the government sector, representing the cost of production, had always been reflected in the National Accounts. However, there was no measure of the quantity of government output and the real growth in its provision. Because of the perceived complexity of directly measuring government output, the convention had been to use the quantity of inputs into government 'production' as measures of the quantity of output.[2] Yet this assumption directly leads to zero productivity growth being imposed on the output of the government sector in the

1 The author is Acting Chief, Macro-Analysis Section in the Economic Studies and Policy Analysis Division of Finance Canada. The opinions expressed in this study are those of the author and do not reflect in any way those of Finance Canada. I would like to thank, without implicating, Anthony Atkinson, David Caplan, Alwyn Pritchard, Benoît Robidoux, Andrew Sharpe and Karen Wilson for helpful comments. All the other usual caveats apply. Email: abIorwerth.Aled@fin.gc.ca.

2 Government output here refers to goods and services provided by the government, such as public administration, health care and education. Expenditures on these types of goods and services provided by government currently account for around 20 per cent of GDP in both Canada and the United Kingdom. Government output does not include transfers to individuals and businesses.

Chart 1
The Pattern of Government Spending in the United Kingdom
1995=100



Source: Pritchard (2003).

National Accounts. Starting in 1998, the ONS had begun to introduce 'non-input-based' measures of government output into the United Kingdom's National Accounts.[3] The purpose of Atkinson's report was to comment on the progress that had been made in the United Kingdom and to suggest potential improvements.

Atkinson produced an Interim Report in July 2004 and a Final Report in January 2005 that laid out general principles (Appendix 1) for introducing measures of government output and provide recommendations for improvement in data and methodology (Atkinson, 2004, 2005). The Final Report largely builds on the Interim Report. The principles and methodologies will be discussed in this article. Atkinson concluded that introducing measures of government output is desirable: "we believe that ONS, like other statistical offices, should work vigorously on this agenda." (Atkinson, 2004:128).

The Atkinson Report also includes detailed chapters on ONS measures of government output by sector and the difficulties involved. For example, the input measures initially developed to cover the health sector had been criticized (because they were more akin to input measures) and the ONS has subsequently introduced changes. The detailed suggestions on the four major spending functions of government — health, education, public order and social protection — discussed in Chapters 8-11 of Atkinson (2005) are not discussed in this article.

The Atkinson Review had been prompted by concern over the results from looking at the 'non-input-based' measures. The United Kingdom Government had recently been increasing expenditures on the provision of government output (health and education in particular). It appears to have been of some concern that the measures of government real output developed by the ONS suggested that public sector productivity growth was weak, if not declining. As seen in Chart 1, there had been an increase in expenditure on government spending of over 35 per cent over 1995-2001, but a productivity decline in the provision of those services of 2 per cent.[4]

The first section of the review article addresses the major issues in measuring government output, including what is government output, problems with the current system, and the

3 The United States had a Federal Measurement Productivity Program which produced detailed estimates of government labour productivity by function for the 1967 to 1994 period. However, these estimates were not integrated into the National Accounts (Fisk and Forte (1997). It would seem that an early report was done in Sweden in 1986 (OECD, 1999a).

4 See Tuke (2004) and Caplan (2006). The *Daily Telegraph* stated that the ONS "parachuted Atkinson in after some figures showed that, despite the extra billions Gordon Brown has pumped into the public sector, its productivity — or economic efficiency — had plunged" (24 July 2004). This newspaper also reported that "leaked minutes from a Cabinet meeting in May 2004, recorded that ministers lamented poor levels of public sector productivity" (6 June 2004).

potential benefits of the measurement of government output. The second section examines methodological difficulties in the measurement of government output. The third section discusses international experience in the direct measurement of government output. The fourth section looks at the impacts of the introduction of direct measures of government output. The fifth and last section concludes.

Issues in Measuring Government Output

The basic objective of the ONS was to introduce direct measures of the quantity of government goods and services into the National Accounts (e.g., number of pupils as a measure of education output, number of solved crimes for police). In this manner, an index of government outputs would be introduced and inter-temporal comparisons of changes in real output could be made.[5] Government output could then be introduced into a measure of the growth in real GDP. With the new method generating estimates of real government output, changes in this output relative to the factors used in producing it could be compared over time to give some idea of productivity growth. Anecdotes suggest, for example, that large-scale technological progress has occurred in the medical field, so an assumption of zero technological change in the health care sector could be seen as unrealistic.[6] A further critical requirement, discussed below, would be to adjust for quality change.

The National Accounts currently reflect the value of the government sector, which includes the wage bill, intermediate purchases of goods and services and depreciation expenses.[7] However, the convention that the quantity of goods and services output produced by government was by definition equal to the quantity of inputs, calculated by dividing the nominal value of inputs by their respective deflators, implied no productivity growth. The absence of market prices or revenue figures for the goods and services produced by the government sector meant that obtaining measures of the volume of government services output could not be done in the same way as for the market sector. The difficulties in directly measuring government output had been thought to be insurmountable. However, the 1993 SNA recommended that statistical agencies introduce direct measures of the volume of government services delivered to households into official GDP-by-industry National Accounts estimates.[8] Eurostat took this recommendation seriously, and in the second half of the 1990s directed member countries to produce direct measures of government output.[9] So the ONS proceeded to try to measure output directly.

What is government output?

First, one should be clear about what is not government output. As a matter of accounting, transfers from government to individuals show up

5 There was no attempt to include a current price measure of output that could be included in the measure of nominal GDP, Atkinson (2005:16). Nominal estimates of the value of government output based on input costs would still be produced, but there would be no relationship between these estimates and real estimates.

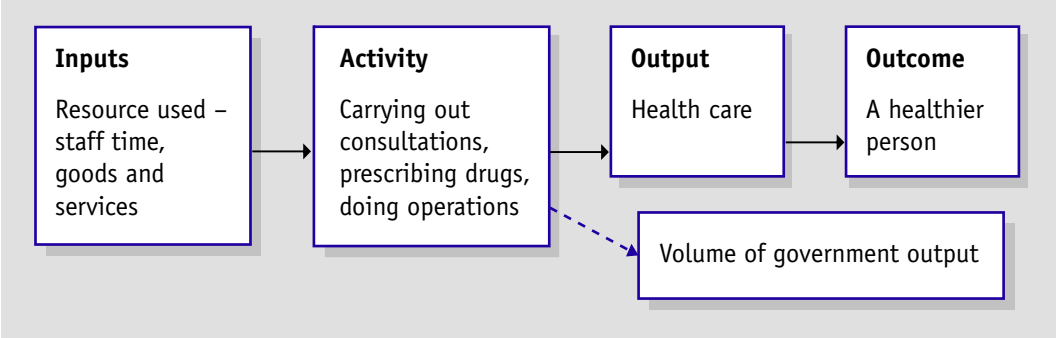
6 The impact of technological change is discussed by Atkinson (2005:18). Indeed, it would be interesting to find out whether the adoption of ICT has had a similar impact in the public sector as it has had in the service sector of some countries since the mid-1990s.

7 Unlike the valuation of market output, no return to capital is included in the sum of costs, but rather only the capital services used up in the production.

8 The *1993 System of National Accounts* (1993 SNA) is a conceptual framework that sets the international statistical standard for the measurement of the market economy. It is published jointly by the United Nations, the Commission of the European Communities, the International Monetary Fund, the Organisation for Economic Co-operation and Development, and the World Bank.

9 Countries (with some derogations) were given a deadline of December 2006 to implement direct measures of government output.

Figure 1
Terminology with Examples Drawn from Health Care



as 'households' income' in the National Accounts. Consequently, 'government goods and services' do not include transfers. The goods and services provided by government show up in the National Accounts as 'government final consumption' (this treatment implying in effect that the government is the final consumer). As an example of how government is treated in the National Accounts, administration of the Canadian Employment Insurance (EI) system is a government service. But the actual payments for EI are not a government service: these payments show up in personal income.

Second, to clarify some of the terminology used in the discussion of government output (without necessarily clarifying how this procedure can be done in practice), Figure 1 shows the path undertaken by government in trying to improve the health of the population. The inputs into the process are the time of medical staff as well as goods and capital services, and the value of these inputs is what is captured currently in the National Accounts. Combined, these inputs undertake certain activities such as prescribing drugs and carrying out operations. These activities constitute health care, the output of government. Conceptually, one can distinguish between whether the item measured is an input (or activity) or an output, but such a clear difference may be difficult to achieve in practice. Input measures may be the number of operations for health care or the hours of lessons for education, whereas output measures may be finished treatment for a disease or the number of students educated.[10]

An output (e.g., health care) leads to the desired outcome of a healthier person. However, individuals can contribute to being healthier through diet and exercise, so the challenge is to capture only the government output that contributes to this outcome: health care does not equate to health. It may very well be desirable to have data on outcomes, but such data may not yield direct information on government productivity.

For a satisfactory measure of government output for the National Accounts, obviously a measure of real output must be found. In addition, a principle of national accounting is to measure outputs: the actual flows of goods and services, and not the final results obtained from their use.[11] Consequently, quantifying the volume of health services through measuring the health of the population would violate National Accounting principles. Atkinson's *Principle B* (Appendix 1) is that the objective should be to measure output from government

10 For example, a proposal to measure education output by the number of hours pupils are taught may be more of an input measure than an output measure.

11 Examples of activities used in Finland by government agencies responsible for the measurement of a particular output are given in Appendix 2.

spending in terms of its incremental contribution to individual or collective welfare, in the same way as market output.

In general, government goods and services can be classified into two types. The first is goods and services that are individually consumed. Within the range of services provided by government, the output and consumers of these goods and services are relatively easier to identify and measure. Examples are health care and education. For some there are private-sector analogues that may facilitate estimating output.

The second type is collective goods and services that are consumed by society as a whole; for example, defense, security and the enforcement of property rights. No one has found a satisfactory method of measuring collective services. For defense, Finland has tried to use the number of conscripts and New Zealand, an index of military preparedness, but there are concerns about the appropriateness of these measures. The output of certain administrative services may be easier to assess, e.g., the number of claims processed. The output of policy advice may be more difficult (although analogous to consulting or legal services in the private sector). Some government services are provided to other government departments, which compounds the problem of measurement.

What are the problems with the current system?

The process of equating the value of output by taking data on input costs can lead to undesirable characteristics in the stated measure of government output (Atkinson, 2005:12 ff.):

- Because the current measure of government output is expressed in money terms (i.e. the dollars expended), changes in the value of goods and services will reflect changes in prices of inputs as well as changes in quantity of inputs and the efficiency by which inputs are converted into outputs. There is no way of isolating efficiency in the use of resources;
- Improvement in the government provision of services could end up as negative output growth under current methods. For example, suppose that the automation of writing cheques displaces clerical workers: the value of labour input is lowered. This labour cost reduction would probably only be slightly offset by increased capital expenditure. Consequently, this change would result in the total cost of inputs declining, which, under the measuring convention, would result in a drop in output. In reality, output would have remained constant or even increased if there were less errors. Increasing substitution of capital for labour in providing government outputs may result in consistent underestimates of government output. This factor may be particularly important in the health care sector, for example.

  More generally, changes to improve efficiency in the provision of government services will not show up in the National Accounts. Under the current convention, there is by definition no change in government productivity. Initial Swedish results had indicated that the cumulative effect of government productivity changes could be quite substantial (OECD, 1999a);
- If government output is an intermediate input into an industry, e.g., through some services provided by government that may not be paid for, then underestimating government output could lead to an overestimate of the value added of the industry; and
- There may be difficulties in interpreting National Accounts data (within a country or across countries) for some goods that can be provided by either the government or the private sector, if the methods of measurement were to differ. By implication, this difficulty carries through to total economy output as well.

What are the potential benefits of measuring government output?

The traditional accounting of government services provision did not reflect how goods and services contribute to living standards. Whether the National Accounts should do so is a point tackled early on in Atkinson's Final Report (Atkinson, 2005:4ff.).[12] Clearly, producing government output involves expenditures, and as such, government outlays should be incorporated in some fashion in the National Accounts. The more difficult question is what are the National Accounts supposed to represent, and given the answer to that question, what is the best way in which government outlays can be incorporated into the accounts. For example, if the purpose of National Accounts is to capture only market activity, then the question of how government output is treated is moot: only government expenditures need be included in the Accounts.[13] Incorporating government output at cost would also be justified if one believes that government is a producer of goods and its technology of production has constant returns to scale and constant technology.

But if one takes a broader view of the National Accounts, then it may be appropriate to include more explicit measures of government output. Although GDP numbers are not a true measure of welfare — GDP does not measure consumer surplus — GDP data are often used as indicators of welfare, despite their numerous pitfalls.[14] Introducing an aggregate measure of government output would certainly be useful information when policymakers choose very broad policy directions. Although it is currently recognized that further investment in health and education is greatly valued to enhance quality of life, discussions of such investments and investments in drivers of economic growth — human and physical capital and innovation — are usually completely separated. There is little data or method to discuss the relative magnitudes in assessing the opportunity cost of various actions at the aggregate level.

There is significant scope for improved information on the overall efficiency and effectiveness in the provision of public services, both within a country and for international comparisons and benchmarking. There is increased concern around the world about improving the quality of public services, in particular to increase government output without necessarily increasing expenditure: in essence, to improve the productivity of providing public services.[15] At the moment, there is no way of measuring overall success in this endeavour. Although incorporating these data into the National Accounts would not be necessary for achieving these goals, going through the discipline of national-accounts methodology would add credibility to any data produced.

Measuring government output could allow improved management of public services by allowing users to (OECD, 1999b):

- Prepare reports on ministries' performance.[16] Note that this implies obtaining

12 Discussion in this section also draws on Yu (2003) and Mamalakis (1996).

13 A possibility advanced in Lynch (2006).

14 The National Accounts try to capture the marginal valuation given by consumers to various goods and services and, as such, measures of the marginal value of government output should be incorporated. Consequently, the challenge for incorporating government output is that the economic meaning of output be consistent with the values estimated for the private sector. The objective is not to introduce a measure of the benefit of government output but rather to measure the incremental effect on outcomes from activity in the public sector. The marginal valuation of these benefits should then be incorporated into GDP: Atkinson (2005:41).

15 For example, see discussion in Danker *et al*. (2006).

16 See Niemi (1999) for Finland.

productivity measures for broad areas such as health care rather than for individual hospitals, for example;
- Monitor benefits of microeconomic reform, as done in Australia;
- Identify areas for reform; and
- Assist in budgetary process.

Equally as important as looking at the results may be the process of getting there. A principle advocated by Atkinson in his Interim Report was that "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 existing indicators)" (Atkinson, 2004:46).[17] Clearly, the preparation of statistics for the National Accounts would require some clear thinking on the part of government departments about their objectives.

Introducing measures of government output would help to improve management of government itself. However, a careful distinction needs to be made between performance measures for individual managers or individual public institutions on the one hand, and the overall performance of government in providing public services on the other.[18] Atkinson (2004) notes the care that needs to be taken in differentiating between output measures and any indicators used to reward performance. Data provided for one type of activity may not be appropriate for the other. Performance indicators need to be simple and precise, but not necessarily consistent over time. On the other hand, data produced for the National Accounts could be transformed in complicated ways, but do need to be consistent over time. Furthermore, there is a danger that an output measure could be manipulated if it became a performance target. For example, assessing performance of schools by the number of pupils that pass an exam may lead to exams being made easier. In that case, the success rate at exams may not then be a good output measure because an increase in this indicator would not reflect an increase in the socially desired outcome. It is therefore important that the data provided on government output for the National Accounts be as comprehensive as possible.

Government reform can make it easier to produce output measures, as has happened in the United Kingdom. Reform usually leads to performance metrics being developed and increased accounting data being produced that could then be the basis for measuring output.[19] Furthermore, establishing more arms-length association between agencies and central government leads to separate accounts being prepared. Experience with performance indicators could be valuable when trying to assess what constitutes government output because it forces people to think carefully about the nature of the goods and services that they are responsible for providing.

Development of performance indicators, when properly implemented, can be useful in providing the imperative to think carefully about the desired output of government. Objective measures of government output could form a base for performance indicators and serve to increase credibility. They may also ensure that performance indicators are more closely aligned with

17 With the United Kingdom having moved on from this point, perhaps this principle was moot for the Final Report. Countries embarking on this endeavour would do well to bear this principle in mind, however.

18 On the use of output measures and their potential impact on performance on university research, see ab Iorwerth (2005).

19 Fisk and Forte (1997) note that the provision of accounting data from federal departments was important to the efforts in the United States of looking at the output and productivity of government programs. Although this information was comprehensive, only government-wide function and summary statistics were published. Nevertheless, these summary tables allowed an overall picture of government productivity to emerge.

the true output of a government department. The objectives that a department sets for itself would have to accord with more tangible evidence that an issue falls within its purview. The intention of Statistics Sweden in 1986 when it introduced direct measures of government output was simply to develop a means of introducing non-zero productivity for government output. However, a debate ensued in Sweden that involved more fundamental questions about the 'value and meaningfulness' of the public services, who was to blame for the results, what had caused the results, and what was going to happen in future (OECD, 1999a). It is not clear that such a discussion should be *prima facie* undesirable.

Incorporating measures of government output could improve international comparability as the practice of directly measuring government output spreads internationally. Traditionally, with zero productivity growth assumed for the government sector, countries with larger government sectors would by definition have lower productivity growth.[20] With some countries moving to non-input based measures of real government output, international comparisons also became problematic. The United Kingdom economy's GDP at constant prices grew at 2¾ per cent a year between 1995 and 2003 whereas the growth rate in the United States was 3¼ per cent. The United Kingdom's approach to measuring government output accounted for nearly half the difference in growth rates (Atkinson, 2005:16).

What are the Methodological Difficulties of Measuring Government Output?

Measuring these outputs is an age-old problem, as many of the same problems apply more generally to quantifying other service industries' output. The standard problems that have arisen in the measurement of government output are that goods and services provided reflect political judgments, outputs are not sold on a market, and the quality of government output.

In the past, the type of services provided by governments may have reflected political judgments, e.g., whether health care should be provided by the government or not. Consequently, measuring government outputs would have reflected political judgments and could have changed with alternating governments. But increasingly in most countries around the world, there is a consensus that at least a significant proportion of health care (for example) should be provided under government auspices. Consequently, measuring government outputs is more likely to be consistent over time.

The discussion in the previous section on the nature of government output was fairly cursory. One could enter into a more profound inquiry into the nature and role of government services, an issue that was not fully discussed in the Atkinson Report. In education and health care, the nature of the good or service may be relatively clear. However, one may have to think very carefully about the service that government supplies in other areas. Much of the transfers provided by government include implicit services analogous to insurance.[21] Employment insurance, health insurance and pensions are all types of insurance services provided by government. But these outputs may not be captured if one only looks at more tangible things such as the operation of the EI system. This problem is analogous to the problem of measuring the true output of the financial services sector.[22]

20 Parenthetically, it is possible that many growth regressions that involve the headline measure of 'government output' as a regressand may be misleading.

21 See discussion and references in Diewert (2001a).

22 For discussion of measuring the role of risk bearing in measuring the output of the financial sector, see Basu *et al.* (2004).

Table 1
ONS Programme of Direct Government Output Measures

Function	Percentage of government spending in 2000	When were output measures introduced	Main components of existing output indicator
Health	30.3	1998	Department of Health Cost-weighted activity index
Education	17.1	1998	Pupil numbers with 0.25% quality adjustment
Social protection — social security administration	2.7	1998	Number of benefit claims for 12 largest benefits
Public order and safety — prisons, courts and probation	3.0	2000	Number of prisoner nights, number of court cases and cost-weighted activity index for probation.
Public order and safety — Fire	1.1	2001	Number of fires attended of different types, other special services.
Social protection — personal social services	7.4	2001	Number of adults and children in care. Number of hours of home help.

Source: Atkinson (2005:15)

Quality

A major difficulty for the effort to measure the real growth in government output is controlling for change in quality. Quality is inherently difficult to measure, even for privately-provided services. In those markets, hedonic methods have been developed to try to incorporate quality changes, such as for the processing capacity of computers. The problem of measuring quality may be exacerbated for government output because of the absence of a market, which may result in excessive reliance on producers to determine quality rather than on consumers' evaluation of it.

A further difficulty posed by quality changes, and also linked to the second issue of there being no market, is that there is no market price. For market goods, quality adjustment is usually done by further deflating the price of the good whose quality has increased. With no market price for government output, such methodologies could not be applied to quality-adjusting government output. Consequently, obtaining direct measures of quality become more important for non-market goods and services because it is only after correcting for quality changes that inter-temporal comparison can be made.

It has been suggested that changes in outcomes could be used to deflate estimates of government output (O'Mahony and Stevens, 2004). The ONS in the United Kingdom is hoping that government departments will voluntarily gather the additional information on quality (that they would need anyhow to monitor their progress) and that these data would be suitable to be incorporated in the National Accounts. Incorporating quality into measures of government output is *Principle B* in Atkinson's report: value should be seen as adjusted for quality.

International Experience in the Direct Measurement of Government Output

What does the United Kingdom do explicitly?

Table 1 summarizes the ONS programme of direct government output measures. The procedure for the measurement of real government output in the United Kingdom is as follows:

1 Produce a list of activities that cover all or most of work done in a particular area (e.g., health care). This practice follows *Principle D*. These activities are not the perfect measure of output but should be reasonably close proxies. The

Table 2
Statistics Reported on United Kingdom Government Output

	1995	1996	1997	1998	1999	2000	2001	% Change 1995–2001
A. Expenditure on government output in current prices*	141,031	146,779	149,147	154,881	166,614	177,801	191,506	35.8
Annual change (per cent)		4.1	1.6	3.8	7.6	6.7	7.7	
B. Volume of government output at 1995 prices*	141,031	142,702	142,779	144,991	149,419	152,524	156,361	10.9
Annual change (per cent)		1.2	0.1	1.5	3.1	2.1	2.5	
Implied deflator* (A_t/B_t)	100	102.9	104.5	106.8	111.5	116.6	122.5	
C. Volume of government input at 1995 prices**	141,031	142,388	141,371	142,785	149,441	153,877	160,320	13.7
Annual change (per cent)		*1.0*	*-0.7*	*1.0*	*4.7*	*3.0*	*4.2*	
D. Annual productivity change: indicative estimate (per cent)** $(B_t/C_t)/ (B_{t-1}/C_{t-1})$		0.22	0.78	0.55	-1.56	-0.86	-1.59	
Implied overall productivity index	*100*	*100.2*	*101.0*	*101.5*	*100.0*	*99.1*	*97.5*	*-2.5*

* National Statistics series.

** Experimental series.

Source: Table 1 and Table 2 in Pritchard (2003).

general approach is that "it is better to measure the right thing approximately than the wrong thing precisely" (Caplan, 1998);

2 For each activity, find a measure of volume, e.g., a count of the number of incidents or cases in a time period;

3 Attach weights to each of the index numbers produced at step 2. The weights used should be proportional to relative expenditure on that activity. Calculate a volume index for the entire area;[23]

To obtain an implied deflator for government output, expenditures on government output can then be divided by the volume of government output. Note that since there are no market prices, this deflator is not a true price index. Accounting for detailed composition of output in undertaking this step becomes particularly important if this input quantity index is used, in conjunction with the output quantity index, to calculate the productivity of government. Atkinson argues in *Principle G* that more detailed coverage should be included and more disaggregated deflators should be used as well.[24]

The processes 1 through 3 can be repeated on the input side in order to obtain a volume measure of government inputs. Subsequently, dividing the volume of output with the volume of inputs yields a measure of productivity.[25] Atkin-

23 Calculating these weights is seen as the principal challenge to further increasing the measurement of government outputs in Finland (OECD, 1999a). But it is also arguable whether the weights should reflect consumer evaluation rather than weights based on expenditures (although this task would be difficult in practice). However, the importance of consumer evaluation is reflected in the activities measured, e.g., crimes solved rather than crimes investigated.

24 Atkinson's concern with using a single quantity index is that it would mask important details and not allow for substitution between various outputs. Consequently, dividing expenditures by a single quantity (such as education expenditure divided by pupils enrolled) may lead to biased results.

25 Note that the interest is in obtaining a measure of productivity and not of technological progress.

Table 3
The Extent of Direct Measurement of Real Government Output in Selected OECD Countries

	Government functions for which volume of government outputs are being directly estimated	Percentage coverage
United Kingdom	Health, education, social protection, public order and safety	Two-thirds of government output
Canada	University education	7 per cent of government output
Finland	Health, education, social protection, public order and safety	70-80 per cent of employee expenditure by central government
New Zealand	Health, education, social protection, public order and safety, public insurance services	Between 60 per cent and 70 per cent of central government output
Norway	Health, education	—
Australia	Health, education	Between 50 per cent and 60 per cent of government output. Experimental measures in other areas.
Netherlands	Health, education	Between 20 per cent and 50 per cent of government output
Italy	Health, education	Between 20 per cent and 50 per cent of government output
Other European Union countries	—	See text

Sources: Various, including OECD (1999a), OECD (2003), Tuke (2004), Atkinson (2004:33), Snelling (2004), Fraumeni *et al*. (2004). Countries may use different methods.

son's *Principle F* states that the measure of inputs should be as comprehensive as possible and include capital services.

Table 2 gives some statistics on United Kingdom government output. Line A gives nominal expenditures on government output. Line B gives a measure of the volume of government output. Dividing line A by line B gives a deflator for government expenditures. A measure of government input volumes is given in line C. Changes in volume of output relative to the volume of input give a measure of annual productivity. Over the 1995-2001 period real government output increased 10.9 per cent compared to 13.7 per cent for government inputs, resulting in a productivity decline of 2.5 per cent.

Atkinson argues, in *Principle H*, that more independent evidence should be advanced to ensure that productivity measures produced by this methodology are corroborated.

What are other countries doing?

Table 3 shows the extent to which direct measures of government output have been introduced in selected OECD countries. At the supranational level, measuring government outputs directly is recommended by the international standard for National Accounts (SNA93, and by the associated European standard ESA95) with coordination taking place at the OECD Public Management meetings.[26] Indeed, all European Union countries under ESA 95, with certain derogations, are committed to measuring government output via the direct volume measure for 2006 data. Consequently the practice of European countries is evolving rapidly as they try to meet this deadline. The latest available developments are given in Malherbe and Gallais (2006a, 2006b). Only 7 per cent of government output in Canada is derived from direct measures of output.

26 Other OECD and Eurostat projects are outlined in Konijn and Gallais (2006).

Table 3 does not include the United States. Health care and education do not always show up as a part of government services in the United States. This difference brings up the issue of where certain services appear in the National Accounts. The services included in the 'government' sector in one country may be in a different sector in the National Accounts of another country. For example, Canada and the United States allocate non-profit institutions (NPIs) to different parts of their accounts, depending on whether government controls or finances the organization. In Canada, NPIs mainly financed by government are allocated to the government sector, regardless of the level of control. Other countries (including the United States) give weight to control. As a hypothetical example, a non-profit hospital financed but not controlled by government is not in the government sector in the United States, but would be in Canada. 'Government'-provided medical care in the United States is a minority of the total medical care provided.[27]

In the GDP-by-industry accounts of the United States, publicly-provided education shows up within the 'general government' sector, mostly in state and local government. Private universities and private schools show up within the 'educational services' sector in the private sector. A large part of education output is part of government output. However, for health care, the majority of output is defined to be in the private sector, although a significant proportion appears under 'state government'.[28] Because of these definitions, the United States' government sector, as defined in the National Accounts, is more limited to government administration and defense.[29] This definitional issue makes the government sector larger in Canada.

These distinctions mean that making international comparisons of the provision of particular services are difficult. However, the difficulties in measurement remain the same since the United States measures the output of these sectors at cost, rather than by direct output measures, wherever they show up. The Bureau of Economic Analysis (BEA) in the United States has been researching improved measures of education and health care.[30] A revised methodology for government transactions has been released (BEA, 2005), but it does not seem to envisage any change to the current input=output methodology.

The Implications of Introducing Direct Government Output Measures

This section gives a limited overview of the implications that introducing direct measures of government output have on the National Accounts.

Aggregate rate of growth of output and productivity

Over 1986-1997, use of the new data has added an average of 0.04 percentage points per year to the growth in real GDP in the United Kingdom. There has been wide variation across years with new data increasing GDP growth by 0.5 percentage points in 1991 but decreasing it by 0.3 in 1995 (Caplan, 1998). There was no systematic impact one way or the other in Italy and Germany, but in Australia and New Zealand GDP growth rates increased with the introduc-

27 In 2003, expenditures for government hospitals was $116.9 billion in the United States, compared to $515.9 billion for all hospitals and $1,557.2 billion for all medical care. See also discussion in Mead *et al.*(2003).

28 I am grateful to Robert McCahill and Brooks Robinson for clarifying some of these points.

29 Lal (2003) states that the relative size for government in Canada and the United States is roughly equal when looking only at administration and defence.

30 See Fraumeni (2004), Fraumeni *et al*. (2004) and Baker *et al*. (2004). Robinson (2006) highlights research that the BEA has conducted, including Christian *et al*. (2006).

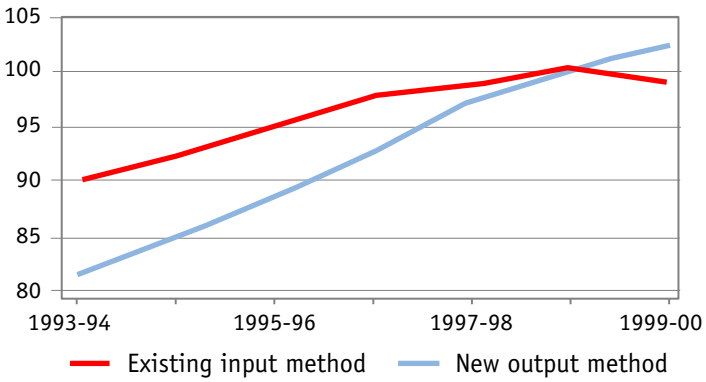
tion of direct measures of government output (OECD, 2003).

At a more detailed level, average annual growth of labour productivity in the UK education sector was 0.86 per cent from 1994 to 2001 (O'Mahony and Stevens, 2004).[31] Economy-wide labour productivity growth in the United Kingdom was 1.44 per cent a year. The same measure of labour productivity for education in the United States showed a decline of 1.17 per cent a year over the same period. Estimates for productivity growth in central government in Finland show a decline of 0.5 per cent in 1995 but increasing by 0.8 per cent the following year and by 1.9 per cent in 1997 (OECD, 1999a).

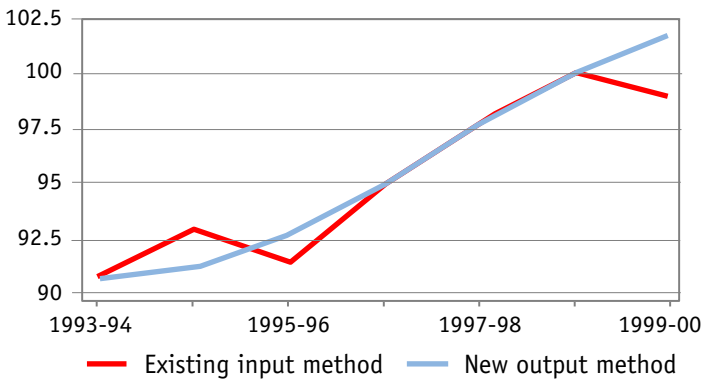
Chart 2 shows the growth path in the chain volume measure of gross value added for health and education in Australia using newly-developed output series and the existing input-based series. The output-based series show an average annual growth of 4.0 per cent for the volume measure of gross value added of health for 1993-94 to 1999-2000 compared to 1.6 per cent under the existing input method. In the education sector, annual growth in gross value added over the same time period averaged 1.9 per cent under the new method against 1.5 per cent under the old method. The output-based series probably capture technological improvements.

Chart 2
Measures of Real Output for Health and Education in Australia
1998-99=100

Health and community services, gross value added
(chain volume measures)



Education services, gross value added



Source: Australia Bureau of Statistics (2001).

The timing of expenditure and output changes

Introducing measures of direct government output are likely to introduce lags into the effect of changes in government expenditures on GDP. The delay would arise because changes in government spending could, for instance, affect the capacity to supply more output rather than increasing output directly (Powell and Pritchard, 2002 and Pritchard, 2003). The spending of money does not necessarily mean that government output would go up immediately. As seen in Chart 2, output-based series are more stable over time. These series capture underlying technological progress and actual output changes rather than changes in expenditures, which may be somewhat more erratic.

Macroeconomic management

Atkinson argues that incorporating direct measures of government output in official real output estimates should not change numbers relevant to

31 Productivity here is defined by an outcome-weighted measure divided by a wage-weighted measure of labour input.

macroeconomic management (Atkinson, 2005:4-5). This position rests on the assumption that the key variable for macroeconomic policy is the output gap; data on private-sector activity are probably more relevant for that calculation. To the extent that government spending affects macroeconomic control, total expenditures may be the most important information because they would give an idea of how much resources are being taken by the government. This position is probably correct in the short term, but as Atkinson notes, improved government productivity could increase overall economic growth in the long run (Atkinson, 2005:5-6). Although not addressed directly in the Report, the lag between the timing of expenditures and the increase in reported output (discussed above) would also suggest that expenditures are more important for macroeconomic policy.

An issue not discussed in the Atkinson Report is the inclusion of the prices of government outputs in a measure of inflation (a CPI). The current practice is to include only prices of activities done in the market, but there is a lengthy discussion in the academic literature of how closely the CPI should reflect a Cost of Living Index (COLI). The current practice is to measure a conditional COLI where the index is conditional on external factors, such as temperature or the level of public goods provision.[32] In principle, a price index developed for public goods could be included in the CPI. It is however unlikely that the public goods price index could be introduced in a timely fashion for such a comprehensive price index. It is also unlikely that there would be consistent changes in a public-goods price index over time.

Cost of introducing a system to measure government output

Government departments in the United Kingdom were already collecting all data used in calculating the new measures of government output because of more general reforms of government that had already been enacted. Consequently the incremental resource costs measures of calculating direct government output were not significant.

Given the extensive structural reforms of government that have occurred in the United Kingdom over the last twenty years, data may be more readily available in the United Kingdom than in other countries. For example, in its reform of the health service, the United Kingdom has established more contractual processes between health providers, which naturally generate a lot of accounting data. It is unclear to what extent these data already exist in other countries: they might exist in isolation, but their collection may not have been centralized.

Conclusion

Anthony Atkinson has done an invaluable service in bringing the arcane world of measuring government output to the attention of a wider audience and in undertaking detailed analysis of the issue. Increased information on government output would be useful in forming a more complete picture of an economy and could be of assistance in performing comparisons across countries, to the extent that countries are moving towards direct measures of government output. Even measuring government outputs approximately may be preferred to no measure at all. However, the most tangible role is likely to be in giving information — at the aggregate level — on the state of public services and on the effect of reform and technological advance on the provision of these services.

Giving a fuller picture of public services is likely to require gathering significant amounts of additional data and introducing accounting systems to reconcile these data. However, it may

32 See discussion in Diewert (2001b), for example.

well be appropriate to ask whether these data and systems should not exist anyhow. Developing data and accounting systems to meet National Accounting standards could be an effective way of ensuring credible and objective data are produced.

An effort to measure government output may have positive implications for thinking about the nature of goods and services provided by government. How do activities undertaken by government departments contribute to increasing government output and thus to the well-being of citizens? What is the nature of the services provided by government, and how can their provision be improved in a tangible and transparent way?

Introducing direct measures of government output in the National Accounts could be done on a piecemeal basis. Given the present concern about the provision of health care, it may be appropriate to develop national indicators of the performance of the health care system as a whole. Such indicators would enable both the public and the government to monitor the impact of any reform or changes in funding.

References

ab Iorwerth, Aled (2005) "Methods of Evaluating University Research Around the World," Working Paper 2005-04, Department of Finance, Canada.

Atkinson, Anthony (2004) *Atkinson Review: Interim Report*, (Norwich, U.K. HMSO) available at http://www.statistics.gov.uk/about/data/methodology/specific/PublicSector/atkinson/default.asp.

Atkinson, Anthony (2005), *Atkinson Review: Final Report Measurement of Government Output and Productivity for the National Accounts* (Norwich, U.K.: HMSO) available at http://www.statistics.gov.uk/about/data/methodology/specific/PublicSector/atkinson/default.asp.

Australia Bureau of Statistics (2001) *National Income, Expenditure and Product March Quarter* (Canberra: Australia).

Baker, Bruce E., Pamela A. Kelly and Brooks B. Robinson (2004) "Estimates of Real Government Consumption Expenditures and Gross Investment by Function for 1959–2003," *Survey, of Current* Business, 84, 10, October, pp. 5-10.

Basu, Susanto John Fernald and Christina Wang (2004) "A General-Equilibrium Asset-Pricing Approach to the Measurement of Nominal and Real Bank Output," paper presented at CRIW Conference on Price Index Concepts and Measurement, Vancouver, June 28-29.

Bureau of Economic Analysis (2005) *Government Transactions Methodology Papers: U.S. National Income and Product Accounts* (Washington, D.C.: Bureau of Economic Analysis) available at http://www.bea.gov/bea/mp5.pdf.

Caplan, David (1998) "Measuring the output of Non-Market Services," *Economic Trends* 539, October, pp. 45-49.

Caplan, David (2006) "Measurement of Non-Market Output in Education and Health," paper prepared for the joint OECD/ONS/Government of Norway workshop, London, England, October 3-5, available at http://www.oecd.org/dataoecd/36/21/37419881.pdf.

Christian, Michael Bruce Baker, Barbara Fraumeni, Alyssa Holdren, and Matthew Williams (2006) "Measurement of Government Output in Education and Health: Alternative Approaches," paper prepared for the 29th General Conference of The International Association for Research in Income and Wealth, available at http://www.iariw.org/papers/2006/christian.pdf#search=%22christian%20iariw%22.

Danker, Tony, Thomas Dohrmann, Nancy Killefer and Lenny Mendonca (2006) "How Can American Government Meet its Productivity Challenge?" McKinsey report, available at http://www.mckinsey.com/aboutus/mckinseynews/pressarchive/pdf/American_govt_meet_prod_challenge.pdf.

Diewert, W. Erwin (2001a) "Productivity Growth and the Role of Government," University of British Columbia Discussion Paper 01-13.

Diewert, W. Erwin (2001b) "The Consumer Price Index and Index Number Purpose," *Journal of Economic and Social Measurement* 27, pp. 167-248.

Fisk, Donald and Darlene Forte (1997) "The Federal Productivity Measurement Program: Final Results," *Monthly Labor Review*, 120, 5, May, pp. 19-28.

Fraumeni, Barbara M. (2004) "BEA's Government Research Agenda," presented at Public Services Performance Workshop, NIESR, London, March 3.

Fraumeni, Barbara M. Marshall B. Reinsdorf, Brooks B. Robinson and Matthew P. Williams, (2004) "Price and Real Output Measures for the Education Function of Government: Exploratory Estimates for Primary & Secondary Education",

paper presented at CRIW Conference on Price Index Concepts and Measurement, Vancouver, June 28-29.

Konijn, Paul and Alain Gallais (2006) "The National Accounts perspective, time and spatial comparisons: the projects of Eurostat and OECD," presented at the OECD/ONS/Government of Norway workshop on Measurement of Non-market Output in Education and Health, London, England, October 3-5, available at http://www.oecd.org/dataoecd/26/30/37426173.pdf.

Lal, Kishori (2003) "Measurement of Output, Value Added, GDP in Canada and the United States: Similarities and Differences," Statistics Canada Research Paper, available at http://www.statcan.ca/english/research/13F0031MIE/13F0031MIE2003010.pdf.

Lynch, Robin (2006) "Measuring Non-Market Output in the National Accounts," presented for the 29th General Conference of the International Association for Research in Income and Wealth, Joensuu, August 20-26.

Malherbe, Francis and Alain Gallais (2006a) "Review of European and OECD Countries' Experiences: Health Output Methods," presented at the OECD/ONS/Government of Norway workshop on Measurement of Non-market Output in Education and Health, London, England, October 3–5, available at http://www.oecd.org/dataoecd/23/3/37443362.ppt.

Malherbe, Francis and Alain Gallais (2006b) "Review of European and OECD Countries' Experiences: Education Output Methods," presented at the OECD/ONS/Government of Norway workshop on Measurement of Non-market Output in Education and Health, London, England, October 3–5, available at http://www.oecd.org/dataoecd/23/4/37443345.ppt.

Mamalakis, Markos J. (1996) "Misuse and Use of National Accounts as a Welfare Indicator: Selected Analytical and Measurement Issues," R*eview of Income and Wealth* 42, 3, September, pp. 293-320.

Mead, Charles Ian, Clinton P. McCully and Marshall B. Reinsdorf (2003) "Income and Outlays of Households and of Nonprofit Institutions Serving Households," *Survey of Current Business*, 83, 4, April, pp. 13-17.

Niemi, Mervi (1999) "Measuring Government Sector Output and Productivity in Finland – Application of the Output Indicator Method," presented at OECD Meeting of National Accounts Experts, Paris, September, 22-25, available at http://www.oecd.org/dataoecd/52/49/2681798.pdf.

OECD (1999a) "Measuring Public Sector Productivity: Case Studies Presented to the PUMA Expert Group Meeting on Measuring Productivity in the Government Sector," presented at OECD Meeting of National Accounts Experts, Paris, 21-24 September, available at http://www.oecd.org/dataoecd/63/1/2681338.pdf.

OECD (1999b) "Productivity Measurement In The General Government Sector," presented at OECD Meeting of National Accounts Experts, Paris, 18-19 March, available at http://www.oecd.org/dataoecd/62/63/2681322.pdf.

OECD (2003) "Measuring the Volume of Government Outputs," presented at OECD National Accounts Expert Meeting, Paris, October 7-10, available at https://www.oecd.org/dataoecd/47/33/24477928.doc.

O'Mahony, Mary and Philip Stevens (2004) "Outcome Based Measures in International Comparisons of Productivity in Public Service Provision: A Review," presented at conference on the Determinants for Performance in Public Organization, Cardiff, May 7.

Powell, Matthew and Alwyn Pritchard (2002) "Measuring Government Output – Mystical or Misunderstood," presented at the 27th General Conference of the International Association for Research in Income and Wealth, Djurhamm, August 18-24.

Pritchard, Alwyn (2003) "Understanding Government Output and Productivity," *Economic Trends* 596, 27-40.

Robinson, Brooks B., (2006) "The Government Divisions's Research Program: 2004-2006 and Beyond," presented at available at 2006 BEA Government Statistics Users Conference, http://www.bea.gov/bea/dn/gov_division_res_agenda.ppt.

Snelling, Janet (2004) "Measuring U.K. Government Output," *The OECD Statistics Newsletter* 20, pp. 4-5.

Tuke, Amanda (2004) "The Atkinson Review of Measurement of Government Output," presented at Conference on Measuring Performance in the Public Sector, London, March 3, available at http://www.niesr.ac.uk/research/aim/docs/tuke-handout.pdf .

Yu, Kam (2003) "Measurement of Government Output: A Review", *mimeo*, Lakehead University, available at http://flash.lakeheadu.ca/%7Ekyu/Papers/GO.pdf.

Appendix 1

Atkinson's Principles for the Direct Measurement of Output from Government Spending

Principle A: the measurement of government non-market output should, as far as possible, follow a procedure parallel to that adopted in the National Accounts for market output.

Principle B: the output of the government sector should in principle be measured in a way that is adjusted for quality, taking account of the attributable incremental contribution of the service to the outcome.

Principle C: account should be taken of the complementarity between public and private output, allowing for the increased real value of public services in an economy with rising real value of public services in an economy with rising real GDP.

Principle D: formal criteria should be set in place for the extension of direct output measurement to new functions of government. Specifically, the conditions for introducing a new directly measured output indicator should be that:

- It covers adequately the full range of services for that functional area,
- It makes appropriate allowance for quality change,
- The effects of its introduction have been tested service by service,
- The context in which it will be published has been fully assessed, in particular the implied productivity estimate, and
- There should be provision for regular statistical review.

Principle E: measures should cover the whole of the United Kingdom; where systems for public service delivery and/or data collection differ across the different countries of the United Kingdom, it is necessary to reflect this variation in the choice of indicators.

Principle F: the measurement of inputs should be as comprehensive as possible, and in particular should include capital services; labour inputs should be compiled using both direct and indirect methods, compared and reconciled.

Principle G: criteria should be established for the quality of pay and price deflators to be applied to the input spending series; they should be sufficiently disaggregated to take account of changes in the mix of inputs; and should reflect full and actual costs.

Principle H: independent corroborative evidence should be sought on government productivity, as part of a process of 'triangulation', recognising the limitations in reducing productivity to a single number.

Principle I: explicit reference should be made to the margins of error surrounding national accounts estimates.

Source: Atkinson (2005).

Appendix 2

Examples of Output Indicators in the Central Government of Finland

Consumer Ombudsman's Office

- Number of petitions to market court,
- Marketing instructions,
- (Number of) contractual terms negotiated,
- Statements on legislative initiatives,
- Cases solved individually,
- Replies to written inquiries.

Courts such as The Supreme Court, Courts of Appeal, District Courts, Provincial Courts and Supreme Administration Court

- Number of cases settled.

Helsinki City Police Department

- The output indicators of public order and security, such as activities directed toward the protection of property and the individual,
- Crime prevention measured by the number of crimes solved,
- The final products of traffic safety,
- Number of permit documents issued (number of passports, identity cards, driving licenses and firearms licenses).

Housing Fund of Finland

- Decisions about loans and interest subsidies measured as the weighted number of decisions.

National Board of Patents and Registration

- Number of patents,
- Number of utility models,
- Number of trademarks,
- Number of pattern rights,
- Company register cases,
- Association register cases,
- Enterprise mortgage cases.

National Food Administration

- Number of letters guiding supervision,
- Number of administrative decisions and memos,
- Number of publications,
- Number of statements,
- Number of training events,
- New instruction materials.

Prison system

- Prisoner-days

Prosecutors' Offices and District Offices

- Number of cases dealt with

State Audit Office

- Number of annual audits,
- Supplementary audits,
- International audits,
- Expertise activities,
- Statements.

Tax Administration

- Numbers of private persons, agricultural entrepreneurs and entrepreneurs and corporations subject to income and property tax,
- Number of supervised registered employers,
- Numbers of primary producers and entrepreneurs subject to value added tax,
- The output indicator of real estate tax

Universities

- Number of degrees completed (generally separated into graduate and postgraduate degrees),
- Adult education and continuing education measured, for example, in days or number of courses (depending on the university),
- Number of publications (research).

Source: Niemi (1999).