

Measurement of Public Sector Output and Productivity

James Douglas

New Zealand Treasury
Policy Perspectives Paper 06/09

November 2006



THE TREASURY
Kaitohutohu Kaupapa Rawa

**NZ TREASURY
POLICY PERSPECTIVES
PAPER 06/09**

Measurement of Public Sector Output and Productivity

MONTH/YEAR

November 2006

AUTHOR

James Douglas
New Zealand Treasury
PO Box 3724
Wellington 6015
NEW ZEALAND

NZ TREASURY

New Zealand Treasury
PO Box 3724
Wellington 6015
NEW ZEALAND

Email information@treasury.govt.nz

Telephone 64-4-472-2733

Website www.treasury.govt.nz

DISCLAIMER

This document was commissioned by the New Zealand Treasury. However, the views, opinions, findings and conclusions or recommendations expressed in it are strictly those of the author(s), and do not necessarily represent and should not be reported as those of the New Zealand Treasury. The New Zealand Treasury takes no responsibility for any errors or omissions in, or for the correctness of, the information contained in this Paper.

ISSN

1176-9513

Summary

The non-market output of government has typically been measured poorly in national accounts, where outputs are often assumed to grow at the same rate as inputs. Statistical standards now recommend that government output should be measured directly. Direct measurement is important both for increasing the quality of national accounts and for measuring government service performance.

Over the past decade, statistical agencies, particularly the United Kingdom's Office of National Statistics (ONS), have invested in improving government output measurement. The United Kingdom's national accounts now measure directly the output of almost 70% of government production, including health, education, administration of justice, social security, fire services, and personal services. The ONS has also made advances in the measurement of output quality change and has used output statistics for public sector productivity measurement.

This paper documents the progress the ONS and other statistical agencies have made in measuring government output for national accounts and for productivity measurement, and considers what this progress implies for New Zealand.

Contents

Summary	i
Contents	ii
Introduction	1
Difficulties with Measurement of Output and Productivity	2
Productivity	5
Output Measurement in New Zealand	6
Progress in Measurement of Output	7
Progress in Measurement of Productivity	11
Adjusting for Quality	11
Education.....	12
Health	13
Implications for New Zealand	14
Conclusions	16
References	17

Measurement of Public Sector Output and Productivity

Introduction

The government is a large, important part of the economy.

The public sector is responsible for producing a large proportion of goods and services in New Zealand. Government production and investment make up over 20% of New Zealand's economy, notably providing the large majority of the health and education services, providing administration of the welfare system and providing justice and defence. The government, like any other part of the economy, produces goods and services, or 'output', by using a set of 'inputs' – labour, intermediate goods and services, and capital equipment. However, despite the considerable importance of the government's output, often little is known about the level of output produced. Often more importance has been placed on how much is spent on inputs in a particular area rather than on how much output is produced as a result.

Traditionally the output of government has been poorly measured.

This predilection for measuring inputs rather than outputs is also reflected in typical national accounts treatments of the public sector. For non-market government services, the measure of output volume in gross domestic product (GDP) is often approximated by measures of input volume such as the number of hours worked (output = input). This output = input treatment implicitly assumes that the amount of output produced per unit of input is not changing, or, put another way, productivity is constant. This approach is deficient, as there is no reason to believe productivity in the public sector should not be changing over time.

However, things are changing ...

In recent years, there has been a push to improve the measurement of output from the public sector to improve the quality of national accounts and to enable measurement of productivity for these services. This has been formalised in official national accounts standards set by the European Union's statistical agency, Eurostat, and the United Nations (Eurostat 1995, 2001, United Nations 1993). Statistical agencies around the world are increasingly encouraged or required by legislation to replace input measures with real output volume measures. To date, considerable progress has been made internationally, particularly in the United Kingdom, and progress has also been made in New Zealand.

This paper summarises the techniques used by statistical agencies to measure the output and productivity of public services and examines the implications for New Zealand of the worldwide drive towards better public service output measurement. The focus of this paper is on national accounts' treatments and uses the United Kingdom as the primary example, because progress is most advanced there. Over time, there will

be increased pressure for New Zealand to update its measurement of the output of the public sector so that our national accounts remain comparable internationally and so that we continue to use best practice. And it will also be important to measure productivity to inform and motivate debate on the performance of the public sector.

Difficulties with measurement of output and productivity

The measurement of output is a crucial part of calculating productivity and assessing performance for any type of organisation. National accountants measure the performance of entire economies in the commonly quoted measure GDP. National accounts give two main measures of the total output of an economy, *current price* GDP and *constant price* GDP. The first gives the total monetary value of the output of the economy each year, using the prices prevailing in that year; the second removes the effects of price changes and represents the actual or *real* volume of output produced in the economy. Constant price, or real volume, measures are typically more interesting from an economic growth perspective because they identify when an economy produces more goods and services. In this paper, when we refer to output, we will mostly be referring to real output. We are also more concerned with growth in output rather than the actual level of output.

For many government services, output growth is assumed equal to input growth...

Typically, the output from public services has been measured poorly in national accounts, where in many cases changes in input levels (often hours worked) are used as a proxy for changes in output levels. This means that the growth rate for the government's *actual* production is not observed and is simply approximated by the growth rate in the number of hours worked. Such a framework is deficient because it does not allow for changes in the amount of output produced per input. This means productivity, the amount of output per unit of input, is assumed to be constant. Labour productivity growth was 2.6% per year during 1988 to 2005 in the well-measured market sector of the economy and there is no reason to believe that productivity growth should be zero in the public sector (Statistics New Zealand 2006).

...which means productivity is constant over time.

Constant productivity assumptions are concerning because of two main reasons. First, growth in GDP is used widely to judge how well an economy is doing over time and in comparison with other economies. Analysing economic growth is particularly important in forming and adjusting economic policy. Constant productivity assumptions for a large proportion of the economy, such as the public sector, may cause understatement or overstatement of growth and could result in misleading conclusions, particularly in cross-country comparisons where the size and structure of public sectors vary widely. The second main reason is that, without real output and productivity measurements, we have no idea how well the government is using its resources. Without productivity growth

measures we cannot tell whether tax payer funds are being put to better or worse use than in the past.

For many government non-market services, the unit of output is difficult to define...

The constant productivity assumption has been made in the past because the measurement of output and productivity in the public sector suffers two prominent difficulties. The first is defining the units of output. For some services, this is easier; for example, an output in health could be the completion of an individual operation or a consultation with a general practitioner. For others, it is more difficult; for example, New Zealand's defence forces contribute to a safer environment for all New Zealanders but it is very difficult to quantify the output of defence. Health is an example of an *individual* service, as the output is delivered to individual consumers, whereas defence is a *collective* service where the output is consumed by the entire economy simultaneously. The relative ease of defining and measuring output in individual services has meant most progress in output measurement has been made for individual services.

Often, individual public services have private sector equivalents and the same output methodology should apply to both. For example, in private sector education, the amount spent on fees would be adjusted for price inflation to get a real volume of output. This would essentially amount to a volume measure based on the number of pupils (potentially with a quality adjustment) and a similar methodology could be applied to public sector education. In the case of government departments providing consultancy services, such as providing advice to ministers, methods used for calculating real output for privately supplied consultancy services could be applied analogously.

...and the value of output is hard to measure without market prices.

The second major difficulty with output measurement is the lack of market price information, due to public sector output being provided free or for an economically insignificant amount. This means that it is difficult to attain a value for each service. In the market sector of the economy, the price of a good or service is taken to represent the marginal value to the consumer; in the non-market environment, we have no such information. This has meant that in *current price* GDP, the total monetary value of public sector output is calculated from the sum of the input costs (thus assuming zero net operating surplus or profit).

The lack of price information means it is difficult to add together diverse units of output. To create measures of aggregate output, we need to sum together the various outputs a service produces. For example, in calculating total health output, we need to somehow combine the number of heart operations with the number of ingrown toenail removals. These two outputs have different values to the consumer and it is misleading to simply add one heart operation to one toenail operation to get a total of two operations. In a market situation, where the value to the consumer is represented by the price, aggregate output is calculated by multiplying the number of operations in each category by the price in each category before adding each type together. This *weights* each output by its value in the total aggregate output. Since public services are usually provided at no or below cost, other means of weighting each output are needed.

The ideal practice would be to weight outputs using a measure that represents the marginal value of the service to the consumer. In health, for example, this could be done using a measure like quality adjusted life years (QALYs), which gives an estimate of the total benefit of an intervention to the patient in terms of added life years. However, comprehensive sets of QALY information are often lacking. An alternative for services such as health or education is to use prices from private health care and education as proxies. This would assume that private education and health operated in competitive markets and that the quality of private and public provision is the same.

In the place of value, cost of production is often used.

In practice, it is difficult to estimate what value each output has and typically, in summing together public service outputs, *cost weights* are used. This means that each different type of output is weighted by the cost of providing that output before the outputs are added together. By doing this, we are using the relative per-unit costs as a proxy for the relative per-unit values to the consumer. In a competitive market, where all output is allocated until marginal cost is equal to marginal value, this may be valid. However, for a public service that might be under or over allocating services, this is not ideal and should be avoided if actual value weights are available.

Outputs often change in quality over time and this must be accounted for.

Measurement of output is also complicated by the changing nature of some outputs over time. Outputs often change in quality and this should be accounted for; for example, heart operations may be augmented by new technology and techniques, resulting in better outcomes for patients. Some quality change can be accounted for by differentiating output into categories that are as homogenous as possible. Doing this means that quality adjustment is recognised in relative volume changes between output categories. When it is no longer practical to further differentiate output, then quality adjustment can be recognised by marking up or down the volume of output according to a quality index. This effectively treats higher quality output as a higher quantity of output. For example, a higher quality heart operation could be worth 1.5 original heart operations.

Quality adjustment is a complex issue and remains problematic (not just in this area but in the construction of price indices, where it is important to distinguish between changes in price due to quality changes and “pure” price changes). Some aspects of the public service are more amenable to quality adjustment than others; for example, QALY measures can be used as an indicator of the quality of health care. For education, the amount of improvement in a cohort’s test results could be used to measure quality change, although this is complicated by the other influences on test results outside the control of the school system. In comparison, there is no obvious form of objective measurement of changes in the quality of policy advice provided to ministers.

Despite the difficulties involved, recently there has been pressure to improve public sector output measurement in the national accounts. The *Atkinson Review* commissioned by the United Kingdom’s national statistician and Eurostat’s *Handbook on Price and Volume Measures in the*

National Accounts provide advice to statistical agencies and set new standards for measuring output in national accounts (Atkinson 2005, Eurostat 2001). Both strongly discourage the use of inputs as a proxy for outputs in calculating output growth from public services in constant price GDP. For individual services, such as health and education, output measures based on input proxies are no longer deemed acceptable by Eurostat. For collective services, where output continues to be very difficult to measure, Eurostat and the Atkinson Review both concede that input measures may still be the most appropriate, provided they are as comprehensive as possible.

EU members must now use direct output measures.

Decisions made by the European Commission mean that European Union members are required from 2006 by legislation to use the standards set by the *Handbook on Price and Volume Measures in the National Accounts* (Decision 98/715/EC and 2002/990/EC). This means that from 2006, European Union members must use comprehensive direct output measures for individual services. The OECD will also launch a project in October 2006 to provide more detailed international guidelines for the development of volume measures of non-market output, in particular for education and health.¹

Productivity

With output and input measures, we can calculate productivity...

The movement toward better output measurement of public services in the national accounts means that productivity has become to a limited extent, and will become increasingly, measurable using national accounts data (as has become routine for the majority of the market sector of the economy; see Black 2003, Statistics New Zealand 2006). The other ingredient needed to calculate productivity is input information. Depending on the productivity statistic required, we need the amount of labour input, the amount of capital input and the amount of input goods and services. The Atkinson Review recommends that, for public sector productivity calculations, the input measures should be as comprehensive as possible covering all three areas.

...however, input measurement requires care too.

All inputs must be measured in real volume terms either by measuring the input directly (ie, recording the number of hours worked) or by deflating expenditure on inputs by an appropriate cost index. The measurement of input from capital is often problematic and is often measured by capital consumption – the depreciation of fixed capital. This is not ideal, as it does not include the opportunity cost of the capital being employed in the public sector, and the Atkinson Review recommends that the appropriate measure is the flow of capital services. This involves adding an interest charge on the entire owned capital to capital consumption. It is also important to consider quality changes in the inputs in productivity calculations, as output

¹
http://www.oecd.org/document/34/0,2340,en_2649_201185_36450978_1_1_1_1,00.html

may change due to different-quality workers, machines or intermediate goods and services being employed. However, as with outputs, making quality adjustment is difficult.

The ONS has recently calculated productivity growth in the public sector by subtracting the growth in *all* inputs, including intermediate goods and services, from the growth in *gross* outputs. Comparing gross output growth with total input growth provides an indication of how the amount of output the public is getting for each dollar of expenditure is changing when the effect of input price movements are removed. This type of measure is different from the type of measure typically used when considering productivity growth in the market sector. For the market sector, *value added* (gross output minus intermediate inputs) is usually compared with labour growth to get labour productivity, and then with both labour and capital input growth to calculate multi-factor productivity. The difference means that public sector productivity growth calculated in this way is not comparable with growth rates calculated for the private sector.

Macro-level productivity indicators do not tell the whole story.

It is also important to note that productivity calculated in this way is a residual of two difficult measurements – inputs and outputs – both of which may include considerable error. Productivity calculated in this way should not be used alone to judge a service’s performance, particularly given the politically sensitive nature of productivity conclusions. These macro-level indicators should feed into more comprehensive productivity studies that use micro-level information and look at changes in outcomes. The Atkinson Review recommended a process of “triangulation” where the productivity = outputs/inputs equation is looked at from multiple angles to build up an overall view of productivity.

Output measurement in New Zealand

New Zealand has followed the standards of the United Nation’s System of National Accounts in measuring the output of the public sector. For the entire non-market government sector, output in *current prices* is measured as the sum of the input costs (state-owned enterprises are considered to be part of the market sector and are measured accordingly). In *constant prices*, most government output measures are calculated using volume extrapolation, where base year value added is extrapolated using a volume index. For government administration and defence, the volume index used is based on deflated wages and salaries, an input measure. For education and health, the volume measure is based on direct output measures (McGrath 1999, Statistics New Zealand 1998).

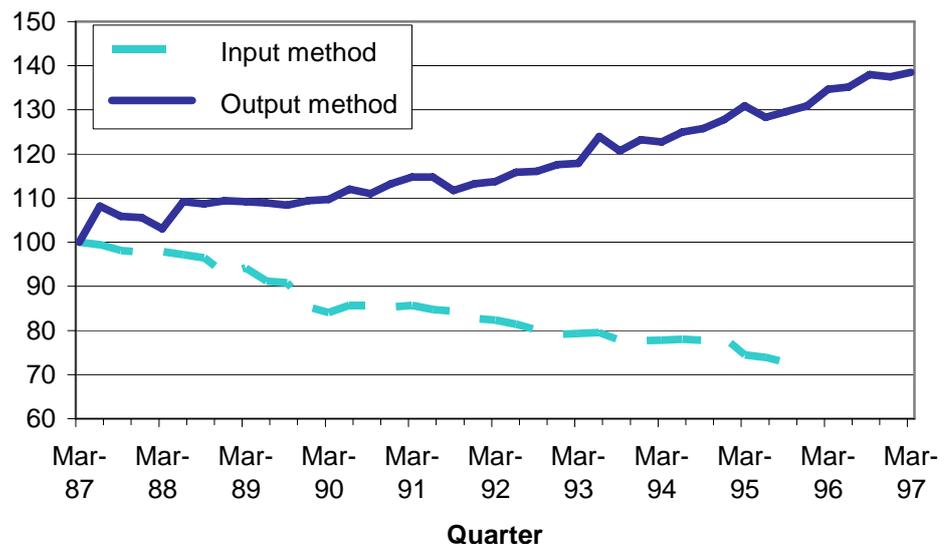
NZ’s national accounts contain output measures for health and education.

For public hospitals, direct output methods were introduced in 1996 to replace a method that was based on a mix of input and output volumes. The output volume measure introduced is based on the number of inpatient discharges adjusted by case mix, the number of day-patient discharges, and the average length of stay in hospitals. Figure 1 shows the considerable difference adopting the new output measures made to the volume of public hospital output in New Zealand’s national accounts. A

volume of output measure for education based on the number of pupils taught at different types of educational institution has also been introduced. Both these output volume measures have been backdated to 1986.

On introduction, these real output volume measures put New Zealand among the world leaders in measuring the output of public services. However, since the introduction of these measures, there has been considerable progress internationally in the measurement of output and in making quality adjustments that could be applied to New Zealand. In particular, the method for public hospitals requires a more comprehensive treatment of day-patient discharges that adjusts for complexity and quality of the treatments. The use of average length of stay should also be examined, as length of stay is not an output of hospitals; public hospitals exist to provide health services, not hotel services. Length of stay may change the quality of treatment and, if this is so, length of stay could be used explicitly as a quality adjustment. Both education and health output require quality adjustment.

Figure 1 - Comparison of input and output methods for public hospitals in New Zealand (seasonally adjusted)



Source: Statistics New Zealand

Progress in measurement of output

The ONS is a world leader in public service output measurement.

The ONS has over the last decade made considerable headway in improving measurement of output from the public sector and is a world leader. In 2005, following the final report from the Atkinson Review, the ONS formed the Centre for the Measurement of Government Activity to take forward the recommendations from the review, driving towards better measurement of government output and productivity. This built on a series of changes implemented during the previous decade.

In the UK, output measures for health, education and social security were adopted in 1998.

The first changes from the use of inputs to proxy outputs in the United Kingdom national accounts occurred in 1998 when new volume measures for health, education, and social security were introduced using output series backdated to 1986 (Caplan 1998). This work followed from the general agreement that the assumption of zero productivity change in the public service was no longer acceptable. Also, public service reform had enabled change by making a better range of output data available.

The health output measures were based on a combination of a cost-weighted index of hospital output for secondary care and an index of tests, prescriptions dispensed, and consultations for primary care. The education volume index was based on the number of pupils at each level of education. This measure was quality adjusted by 0.25% per annum based on trend improvements in the results of the General Certificate of Secondary Education exams taken by secondary school students aged 14 to 16. This adjustment assumed that exam outcomes were improving because of general improvement in the quality of education. For social security services, the number of claims for different types of benefits was used, the different types being weighted together using the cost of administering each type of claim. Introducing these output measures added, on average, 0.04% each year to the growth rate in the United Kingdom's GDP during the 1986 to 1997 period.

Output for administration of justice was added in 2000.

During 2000, the ONS introduced new output measures for administration of justice and agricultural intervention in the national accounts (Baxter 2000). The output measures were cost-weighted series backdated to 1994 for prisons, legal aid, Crown prosecution service, and courts. Prisons output was based on the annual average prison population. Legal aid, Crown prosecutions service, and courts output were based on the number of cases handled by classification. The agricultural intervention output was based on the Intervention Board's own cost-weighted activity index. These changes resulted in a 0.07% increase in the level of GDP for 1999.

Output measures for the fire service, probation services and personal services then followed in 2001.

The ONS continued to improve in 2001, adding further government output measures to the national accounts for the fire service, local authority personal services and probation (Ashaye 2001). Direct output measurement in the United Kingdom now covered just less than 70% of total government spending on output. The new output measure for the fire service was a weighted index of the number of fires attended and false alarms, the number of hours spent on fire-prevention activities, and data on road accidents and other special services. The output for personal social services was a cost-weighted index of the number of children in different types of care, the number of elderly people in homes and the number of contact hours for home help. Probation was based on the probation service's own cost-weighted activity index. The introduction of these new output series reduced the implied growth rates for these services and, over the 1995 to 2000 period, reduced GDP growth by 0.1%.

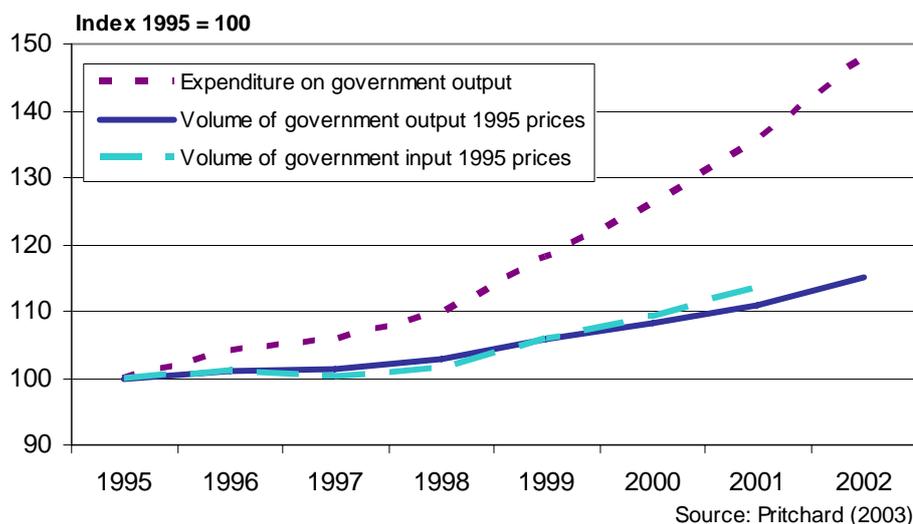
Figure 2 shows an index for the output of the United Kingdom government up to 2002 implied by the output measures described above. The implication of the measures was that output had grown much less than

expenditure on output and also less than inputs. This prompted concern and criticism that the output measures were not fully capturing increases in government output. Following this, the United Kingdom's government statistician commissioned the independent Atkinson Review into the measurement of government output and productivity.

The Atkinson Review began in 2003 and independently reviewed the ONS's use of output measures.

The ONS worked closely with the Atkinson Review team and in 2004, on interim recommendations, used improved output data from the National Health Service in the national accounts (Pritchard 2004a, b). The main changes were to break down treatment categories into smaller, more homogenous categories, changing from 16 to around 1700 categories, and to increase the coverage of the output measures. This meant that the output measure took much greater recognition of trends in the use of different types of treatments at a fine detail and were better able to capture changes from low-quality treatments to higher-quality ones.

Figure 2 - United Kingdom government output, expenditure on output, and government input



The Atkinson Review at its conclusion in 2005 made a number of recommendations for improving the ONS's methods and, soon after the review, a number of changes were implemented. Changes were made to the volume output measures for health, education, personal social services, administration of social security, and fire services (UK Centre for the Measurement of Government Activity 2005). The changes made were incremental improvements in methodology and resulted in a small 0.5 percentage point increase in cumulative GDP growth from 1995 to 2003. A number of recommendations from the Atkinson Review remain outstanding, notably the introduction to the national accounts of comprehensive quality adjustment for public services.

The ABS has introduced output measures for health and education.

Along with the ONS, the Australian Bureau of Statistics (ABS) has also investigated output volume measures in a number of service areas. In Australia's June 2001 national accounts, volume measures for health and education (backdated to 1993/94) were introduced to replace the previous input = output methodology (Australian Bureau of Statistics 2001). For

health, the ABS uses a cost-weighted index of volumes in 660 different treatment categories to measure hospital output and cost-weighted indices for nursing homes, specialists, general practitioners, and pathologists. The new output-based method increased health output growth by a considerable 2.4 percentage points per annum. The volume estimates for education were mainly based on annual student numbers, with a component for university research output based on the number of publications and student research project completions. The new education measure increased education output growth by 0.4 percentage points per annum. Neither of these measures captures quality change aside from compositional quality changes.

The ABS investigated output measures in the justice sector but lacked good data for police.

In 2001, the ABS produced a discussion paper on output volume measures for police, justice services, and corrective centres (Northwood, Hinchcliffe, Henderson and Rawnsley 2001). This paper investigated the way output should be measured in the justice sector and constructed experimental indices for each part. The output of justice services was measured by the number of cases finalised, broken down by court level and jurisdiction. The output of corrective centres was measured by the number of prisoner days in various kinds of detention programmes. The measurement of output of police was formed based on estimates of the number of investigations completed, but this was confounded by lack of appropriate data. Because of this difficulty deriving an index for the police, it was decided that input-based measures should remain for the justice sector, pending better data for police output.

In 2003, the ABS produced unpublished estimates of volume of output measures for the Australian Taxation Office and the Australian Government social security agency, Centrelink (Australian Bureau of Statistics 2004). The new measures aimed to measure the volume of services the organisations deliver to government, businesses and the community. Initial results were promising but the data time series was short and it was concluded that the behaviour of the data should be examined over a longer period before the output measures could be added to the national accounts.

The US uses input = output methods but is looking to change to direct output measures.

Other countries have also experimented with output measurement of the public sector for their national accounts, although no major economies have treatments as comprehensive as those of the United Kingdom. The United States uses input = output methodology for the public sector, despite a long history of productivity measurement. Because health care in the United States is mostly produced by the private sector, the main focus in health output measurement has been on the correct measurement of price changes. In education, the United States government has a larger role and the Bureau of Economic Analysis is moving toward real output measures of the education function in the United States national accounts (Christian and Fraumeni 2005). This is part of a larger research project to measure real output of federal, state, and local governments (Bureau of Economic Analysis 2006). Italy and the Netherlands have also used volume of output measures in their national accounts (Collesi 2000, Kazemier 1997).

Progress in measurement of productivity

Following the introduction of direct output measures in the national accounts, the ONS set about measuring public service productivity (Pritchard 2001, 2002, 2003). The productivity measure they used compared growth in gross output (not value added) with growth in all inputs – labour, capital, and intermediate goods and services. Gross output was chosen, as public and political interest is usually in the total output of each government service rather than the output net of intermediate goods and services. This becomes a particularly important distinction when government contracts out a large proportion of a particular service.

The output measurement used by the ONS in the productivity calculations was the constant price output measure for each service from the national accounts. The input measures were based on staff numbers or pay, the procurement of goods and services required for production, and the depreciation of plant and equipment used in production (capital consumption). The amount spent on each input was deflated by an appropriate price index to get a real volume measure.

In 2003, using its output measures, the ONS found government productivity had decreased.

In the 2003 paper, Pritchard found that the volume of government inputs went up by 14% between 1995 and 2001, while outputs grew by only 11%, implying a reduction in productivity (see Figure 2). Pritchard acknowledged that this negative productivity growth could have been due to increased spending on things that would enhance future capacity, incomplete coverage of output measures, or spending on increases in quality or equality.

The measured productivity of the National Health Service was also found to be decreasing.

Following from this investigation into general government productivity, the ONS focused on a significant area of government spending – health. The ONS investigated productivity in the health service using improved output series from the national accounts and experimental input series (Lee 2004). Between 1995 and 2003, outputs were found to have grown by 28% while inputs had grown by between 32% and 39%, depending on the measure used, implying a productivity decrease of 3% to 8% over the nine years. Hemingway (2004) provides more detailed information on the sources and methods used in Lee's study.

Adjusting for quality

In the ONS's first productivity measures, there was little treatment of quality improvement.

The ONS's productivity measurements for general government and for health resulted in controversy, which was part of the motivation for the Atkinson Review. This controversy was in part due to disagreement about how output should be measured, where a major criticism concerned quality adjustment. Following the Atkinson Review, the ONS has made considerable attempts, in cooperation with spending departments and using expert advice, to improve the measurement of quality of output.

Education

This occurred first in a productivity analysis for education that presented both experimental input treatments and also experimental output series with new treatments of quality (UK Centre for the Measurement of Government Activity 2006a). The analysis aimed to improve on the quality measure used in the United Kingdom national accounts that was based on the past trend in General Certificate of Secondary Education (GCSE) exam results. This trend adjustment had not been updated since the measure was first introduced in 1998.

The ONS looked at new quality adjustment methods for education.

To remedy this situation, the ONS tested a number of different potential adjustments for quality and for the increasing value of education, adjustments that were based on suggestions from the Department for Education and Skills (2005). The first simply looked at more recent trends in GCSE results data. The second used progress between the Key Stages of the English compulsory education system, where a cohort's Key Stage test results were compared with their results at the previous Key Stage (Key Stage examinations are usually taken at age 7, 11, 14 and 16). The third involved adjusting output to account for the increasing value of education as the economy becomes more productive, the idea being that as real wages increase, education provides more benefit to the student. To do this, the ONS made an adjustment based on the trend in real wages in the economy. The ONS discussed potentially using teaching assessment or class size as a measure of quality but used neither. Teacher assessment grades were found to be unsuitable because of frequent changes in the assessment method and the effect of class size on students' performance is heavily debated in the international literature.

Adopting different adjustment methods gave a wide range of productivity growth rates.

To analyse productivity growth, a number of different methods of calculating the amount of input into education services were investigated, including changing the treatment of the cost of labour and using capital services instead of consumption. The total input measure was relatively insensitive to the method used. Combining the various input measures with the various quality and value adjusted output measures created a wide range of potential productivity growth rates from -2% per year to 2% per year over the 1998 to 2004 period, highlighting the significant differences alternative methods create.

It must be noted that the quality adjustments the ONS considered were based on the outcomes of the education system; namely, test results. Using outcomes, like test results, can be misleading due to the multiple contributors to outcomes, although this problem is reduced when a cohort's improvement in standardised test is used. Ideally, the contribution from education service would be separated out from other inputs using regression analysis. However, the data requirements for such analyses are often prohibitive. Adjusting for the increasing value of education in a growing economy is more part of a human capital approach to examining the value of education output. A comprehensive human capital approach would examine the difference in lifetime earnings of people with different levels of education to obtain a value for education. Once again, data

requirements for such a treatment are large and would also involve significant lags. The use of regression analysis and human capital are discussed by Christian and Fraumeni (2005).

Health

New methods of quality adjustment in health were investigated.

A major criticism of the article on health productivity by Lee was that the measures did not contain adjustments for changes in quality. The United Kingdom Centre for the Measurement of Government Activity (2006b) responded to this by producing new measures with various adjustments for quality. The Atkinson review also recommended that health care should be attributed a higher value as the economy becomes more prosperous and healthy individuals can earn higher real wages than previously. In this work, the ONS presented a wide range of productivity estimates, starting from those implied by the national accounts output measures unadjusted for quality, and using a range of different input series. These estimates implied that productivity decreased by between 0.6% and 1.3% per year over the 1995 to 2004 period.

The new adjustments changed the productivity growth rate from negative to positive.

The output measures from the national accounts were then adjusted for quality using a number of indicators including: survival rates, measured increases in patient health due to intervention, life expectancy, waiting times, improved outcomes from primary medical care, patient experience, and the use of value weights instead of cost weights for drugs used to control cholesterol (statins). Combining the quality adjusted measure with the input measures gave a range of 0.2% to -0.5% per annum productivity growth over the period from 1999 to 2004. Following this, an adjustment was made for the increasing value of health due to rising real wages. The third set of estimates indicated productivity growth of between 0.9% and 1.6% per year. This work demonstrates the considerable effect adjusting for quality and value can have, changing the measured productivity growth during the period from negative to positive.

The work on quality of health output was based on research carried out by the Department of Health and by the Centre for Health Economics at the University of York and the National Institute of Economic and Social Research (Department of Health 2005, University of York and National Institute for Economic and Social Research 2005).

In this improvement of health output measurement, the use of value weights instead of cost weights for statins presents an interesting example. The Atkinson Review recommended that the ONS and the Department of Health should identify treatments where marginal valuation and cost weights were likely to be very different. The Department of Health looked at statin use in the treatment of heart disease. Instead of cost weights for these drugs, the Department of Health used value weights based on added life years as a direct result of statin therapy. The value of taking the drug was calculated from estimates of the number of life years gained by patients taking the drugs. In this case, the value weight was over four times greater than the unit cost of the drug, increasing the overall output of the National Health Service.

Both the education and health articles presented a wide range of productivity estimates that made conclusions about performance difficult. In fact, following the publication of the health article, the ONS drew criticism from the Economist (Economist 2006) for publishing “confusing measures of NHS productivity”. However, the criticism seems unwarranted given that these articles were published as ongoing research into the measurement of National Health Service productivity, and the range of measures and methods were published to create discussion and public debate and stimulate work in the field. The ONS recognises that it will take a good deal of discussion and debate before a definitive measure of quality change can be accepted and applied as a standard. The experience of the ONS highlights the length of time required to build comprehensive and fully representative measures of output.

Implications for New Zealand

New Zealand has not made a great deal of progress in introducing new measures of government output since the introduction of the real output measures for education and health in the 1990s. Including the public sector in official measures of productivity is part of the long-term agenda for Statistics New Zealand, but currently this remains a “distant goal” (Statistics New Zealand 2006).

New Zealand needs to investigate wider use of government output measures.

Progress on this front is needed in the near future to ensure the continuing comparability of New Zealand’s national accounts data with other members of the OECD. This is particularly pertinent given the European Union decision to require members to produce output measures for individual government services from 2006 and given the amount of time it takes to develop robust measures. Public service output measures are also crucial for the evaluation of the performance of these services. Robust output and productivity measures from Statistics New Zealand would assist in providing evidence of how government services are performing, although such measures will not provide conclusive evaluations of performance.

There are a number of individual services where progress could be made by applying the methodologies used in the United Kingdom. For example, a survey of fire service annual reports shows that the measures used in the United Kingdom national accounts may be available here in New Zealand. Similarly, the New Zealand Police record statistics on the number of cases investigated in various categories with rates of resolution, which could provide a quality adjusted measure of output for part of police expenditure. Data is also collected on road policing, including data such as the number of fines issued and breath tests administered, which could be included in output measures. However, aspects such as crime prevention by police services will remain difficult to measure, as will the output of our defence forces. Other justice system output indicators, including the number of prisoners and people held in remand, and the number of cases heard at different levels of the courts, should also be considered for the national accounts. On the welfare side, measures of the number of claims for

various benefits and the number of interventions made could be considered.

New Zealand should be able to learn quickly from overseas research.

Once true output measures for the public sector replace input measures, then measuring quality improvements is the next important step. Education and health are the government's biggest spending areas and adjustment for quality could make significant differences to the measures of output for these sectors, as it has in the United Kingdom. New Zealand could experiment with using the quality adjustment methods the ONS has used and is in the position to learn quickly from the United Kingdom's recent research.

At this stage of development, significant problems are likely to occur with the amount of data available to be used for output volume measures and quality adjustment. Data sets often take a considerable period to set up and collection of data needs to be done over a reasonable time period to be useful for national accounts and for analysis of changes in productivity. It is therefore vitally important to plan early so that if appropriate data is not available, steps can be taken to ensure future availability. The experience of the Review of Government Services in Australia shows that it is important not to be put off by lack of data quality, as the publication of provisional data often encourages improvement of data quality and timeliness, resulting in a "journey of continuous improvement" (Banks 2002). Australia's government services review started with very limited data but, over successive publications, data availability significantly improved. By placing great importance on the development of indicator frameworks and then, over time, filling the gaps in the data collections, a "more accurate and complete picture" of the performance of public services can be developed.

Progress will also depend crucially on cooperation between service departments and statisticians. Measurement of output and productivity is most useful when service providers have contributed to developing measures they feel are appropriate. The building of solid relationships is very important and will take time. Ownership of the process by all involved will be essential for the production and development of meaningful statistics.

It will be important to consider the political sensitivity of producing productivity measures. Measurement of output by itself is not overly sensitive, but when outputs are compared with inputs, these statistics become politically charged. A great deal of care will be required in developing robust and comprehensive measures of input and output for government to avoid producing bad measures that could damage the reputation of national statistics. Any productivity measures that are produced must not be used alone but in concert with outcome measures and micro-level indicators to build up a comprehensive productivity story.

Conclusions

Improving New Zealand's measures of public sector output should become a priority over the next few years. This is necessary to maintain internationally acceptable practices in our national accounts, and is a crucial part of the drive to focus on outputs and outcomes in government service performance measurement. The United Kingdom has made considerable advances in output measurement and New Zealand can use this progress as a guide. Australia also continues to test new methods and the European Union and the United States will in the coming years push forward in the area, providing valuable lessons for New Zealand.

References

- Ashaye, Timi (2001) "Recent developments in the measurement of general government output." *Economic Trends* 576: 41-44.
- Atkinson, Tony (2005) *Atkinson review: Final report - measurement of government output and productivity for the national accounts*. (London: Palgrave MacMillan).
- Australian Bureau of Statistics (2001) "New chain volume estimates for the services sector." Australian National Accounts: National Income, Expenditure and Product, March 2001.
<<http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/5206.0Mar%202001?OpenDocument>>
- Australian Bureau of Statistics (2004) "Measurement of services: Recent ABS experience." Conference of European Statisticians, Economic and Social Council, United Nations.
<<http://www.unece.org/stats/documents/2004.06.ces.htm>>
- Banks, Gary (2002) "Reviewing the service performance of Australian governments." Chairman's Speeches, Productivity Commission.
<<http://www.pc.gov.au/speeches/cs20020220/index.html>>
- Baxter, Michael (2000) "Developments in the measurement of general government output." *Economic Trends* 562: 3-5.
- Black, Melleny, Melody Guy and Nathan McLellan (2003) "Productivity in New Zealand 1988 to 2002." Wellington, New Zealand Treasury Working Paper, 03/06.
- Bureau of Economic Analysis (2006) "Final strategic plan for fy 2006–fy 2010."
<<http://www.bea.gov/bea/about/Director.htm#Strategic>>
- Caplan, David (1998) "Measuring the output of non-market services." *Economic Trends* 539: 45-49.
- Christian, Michael S and Barbara M Fraumeni (2005) "Measuring the education function of government." Bureau of Economic Analysis.
<<http://www.bea.gov/bea/about/background/ChristianFraumeni050605.pdf>>
- Collesi, Daniela (2000) "Volume measures and productivity analysis for the non-market sector: The Italian experience." Istituto Nazionale di Statistica.
<<http://www4.statcan.ca/english/voorborg/2000%20madrid/papers/2000-044.pdf>>
- Department for Education and Skills (2005) "Measuring government education output in the national accounts."
<<http://www.dfes.gov.uk/research/data/uploadfiles/RW45.pdf>>
- Department of Health (2005) "Healthcare output and productivity: Accounting for quality change."
<http://www.dh.gov.uk/PublicationsAndStatistics/Publications/PublicationsPolicyAndGuidance/PublicationsPolicyAndGuidanceArticle/fs/en?CONTENT_ID=4124266&chk=5QmbY7>

- Economist (2006) "Take your pick." 378(8467): 53.
- Eurostat (1995) *European system of accounts*.
- Eurostat (2001) "Handbook on price and volume measures in national accounts."
- Hemingway, James (2004) "Sources and methods for public service productivity: Health." *Economic Trends* 613: 82-90.
- Kazemier, Brugt (1997) "Volume measurement of government output; the Dutch practice since revision 1987." Statistics Netherlands.
<<http://www.cbs.nl/nl-NL/menu/themas/macro-economie-financiele-instellingen/nationale-rekeningen/methoden/occasional-papers/1990-1999/1997-086-p30-pub.htm>>
- Lee, Phillip (2004) "Public service productivity: Health." *Economic Trends* 613: 38-59.
- McGrath, Mike (1999) "Non-market services - New Zealand experience in measurement." Statistics New Zealand.
<<http://www.stats.govt.nz/about-us/events/voorburg-group-services-stats.htm>>
- Northwood, Kristen, Christopher Hinchcliffe, Leigh Henderson and Terry Rawnsley (2001) "Experimental output measures for the Australian justice sector." Australian Bureau of Statistics.
- Pritchard, Alwyn (2001) "Measuring productivity in the provision of public services." *Economic Trends* 570: 67-68.
- Pritchard, Alwyn (2002) "Measuring productivity change in the provision of public services." *Economic Trends* 582: 20-32.
- Pritchard, Alwyn (2003) "Understanding government output and productivity." *Economic Trends* 596: 27-39.
- Pritchard, Alwyn (2004a) "Measuring government health services output in the UK national accounts: The new methodology and further analysis." *Economic Trends* 613: 69-81.
- Pritchard, Alwyn (2004b) "Revisions resulting from improving the methodology for measuring government healthcare output in the national accounts."
<<http://www.statistics.gov.uk/CCI/article.asp?ID=911>>
- Statistics New Zealand (1998) "Measuring output of non-market hospital services." National Accounts Division, Statistics New Zealand.
- Statistics New Zealand (2006) "Productivity information paper (1988–2005)." <<http://www.stats.govt.nz/products-and-services/info-releases/productivity-statistics.htm>>
- UK Centre for the Measurement of Government Activity (2005) "Improvements in the methodology for measuring government output." *Economic Trends* 622: 35-45.
- UK Centre for the Measurement of Government Activity (2006a) "Public service productivity: Education." *Economic Trends* 626: 13-37.
- UK Centre for the Measurement of Government Activity (2006b) "Public service productivity: Health."

United Nations (1993) *System of national accounts*.

University of York and National Institute for Economic and Social Research (2005) "Developing new approaches to measuring NHS outputs and productivity." Centre for Health Economics, University of York and National Institute for Economic and Social Research.

<<http://www.niesr.ac.uk/pubs/searchdetail.php?PublicationID=905>>