THE CENTRE FOR MARKET AND PUBLIC ORGANISATION

The Centre for Market and Public Organisation (CMPO) is a leading research centre, combining expertise in economics, geography and law. Our objective is to study the intersection between the public and private sectors of the economy, and in particular to understand the right way to organise and deliver public services. The Centre aims to develop research, contribute to the public debate and inform policy-making.

CMPO, now an ESRC Research Centre was established in 1998 with two large grants from The Leverhulme Trust. In 2004 we were awarded ESRC Research Centre status, and CMPO now combines core funding from both the ESRC and the Trust.

Centre for Market and Public Organisation
Bristol Institute of Public Affairs
University of Bristol
2 Priory Road
Bristol BS8 1TX

Tel: (0117) 33 10799
Fax: (0117) 33 10705
E-mail: cmpo-office@bristol.ac.uk

Productivity in public services

Helen Simpson

December 2006

Working Paper No. 07/164
Published in Journal of Economic Surveys 23 (2) pp.250-276 (2008)

ISSN 1473-625X

The Leverhulme Trust

CMPO Working Paper Series No. 07/164

Productivity in Public Services

Helen Simpson[1]

[1] *Institute for Fiscal Studies, CMPO and Nuffield College, University of Oxford*

December 2006

Abstract

This paper discusses issues arising in the measurement of productivity in public services. Compared to measuring productivity in the private sector difficulties arise because the output of public services is often un-priced and because some public services are consumed collectively. A key problem is measuring the full range of outputs and quality improvements delivered by public sector organisations that are valued by society. Without comprehensive measures of output productivity statistics may be misleading. I outline methods used in the measurement of both private and public sector productivity and discuss the measurement of public sector productivity in practice. Finally I discuss studies that examine factors underlying productivity differences and productivity growth in public and private sector organisations. Public sector reforms and the use of pilot schemes in public sector organisations present opportunities for research to identify causal effects on productivity.

Keywords: Productivity; Public enterprises

JEL Classification: D24; H4; L32

Acknowledgements

I would like to thank the CMPO for financial support for this paper. I thank Ian Crawford, Carl Emmerson, Rachel Griffith, Jonathan Haskel, Alex Klemm and Stephen Redding who have all helped my understanding of various points discussed in this paper. All errors are my responsibility

Address for Correspondence
CMPO, Bristol Institute of Public Affairs
University of Bristol
2 Priory Road
Bristol
BS8 1TX
hsimpson@ifs.org.uk
www.bris.ac.uk/Depts/CMPO/

CMPO is jointly funded by the Leverhulme Trust and the ESRC

1 Introduction

The UK government has set itself a productivity challenge. Two of HM Treasury's main objectives are to increase the productivity of the economy, (increasing the rate of productivity growth and narrowing the productivity gap with other major economies), and to improve the quality and cost-effectiveness of public services. To assess whether these targets are met requires objective measures of productivity and cost-effectiveness, and to deliver improvements and design policy, understanding what underlies differences in productivity across providers and what drives productivity growth is crucial.

Before proceeding, it is worth noting that while the productivity of public sector services is one factor in determining how effectively public money is being spent, it is not the sole consideration. Measuring productivity differs from a cost-benefit analysis, which might be used to assess the 'value for money' of a new government programme. While measuring the productivity of public services is certainly of interest, society may prefer the public sector to deliver more services or improvements in the quality of service even at the expense of a decrease in productivity. Equally, an increase in productivity may not be welcome if it came at the expense of a decrease in the output of public services. However falling productivity unaccompanied by any increase in output might raise concerns, as might significant discrepancies in productivity across providers in the same sector.

In this paper I outline the issues arising in the measurement of public sector productivity, focusing on the measurement of output of public services. I then discuss techniques that are used in the measurement of productivity for both the private and public sectors. Finally I assess some of the evidence on the factors underlying productivity and performance in the public sector and compare these with research examining the drivers of private sector productivity.

Measuring productivity amounts to relating a volume measure of the output of an organisation or sector of the economy, (goods produced and services provided), to a volume measure of the inputs used to deliver them (employees, equipment, and intermediate goods and services). Much of the discussion around the measurement of productivity in public services focuses on the accurate measurement of outputs.

For some public sector organisations, for example hospitals and schools, information such as the number of consultations with doctors or the number of lessons taught is collected and can be used as a measure of the volume of services delivered. However these counts of activities do not necessarily coincide with the full set of actual outputs delivered by providers; the information available simply may not be comprehensive enough to reflect accurately all of the outputs of public services that are actually valued by society. Moreover, using measures of activities to proxy output may make it very difficult to account for improvements in quality and to capture increases in output when technological developments or changes in the mode of delivery reduce the number of activities required to provide the same service. In other cases outputs are simply very hard to measure, for example in the case of collective services such as defence and the prevention of crime.

Even with accurate measures of the various outputs, in the public sector the absence of information on prices makes aggregating outputs for a single provider or for a sector-level service such as education problematic. When measuring output in the private sector, prices, which in well functioning markets reflect relative marginal valuations and marginal costs of production of different goods and services, serve as weights in aggregation.

The fact that measures of output for public sector services differ from those typically available for market sector firms has in some cases lead to the application of different productivity measurement techniques, for example when comparing productivity across a group of service providers. These techniques do not require prices or specified weights to aggregate outputs, but come with some of their own drawbacks in terms of reliability. Other studies of the productivity of public sector organisations have used partial, or single output measures, for example survival rates of heart attack patients. While these can provide reliable, quality-adjusted measures of particular outputs they may potentially be misleading if used to represent output or productivity at the level of an organisation as a whole, particularly if they correspond to an output being targeted under a performance indicator.

Partial output measures for public service providers have been used in some studies looking at the effects of competition or ownership on productivity. These studies ask whether hospitals or schools that face stronger competition from other providers typically display higher productivity, and whether a causal link can be established between competition and productivity performance. Empirical studies looking at the effects of competition on private sector productivity tend to find that greater competition in product markets is associated with

increased innovation and productivity growth. Obtaining rigorous empirical evidence on whether such findings translate across to the public sector is extremely valuable for policy formation and assessing public sector reforms.

The paper is structured as follows. The next section outlines issues in the measurement of outputs and inputs for use in the construction of measures of productivity. Section 3 discusses different productivity measurement techniques, and how the techniques used to measure productivity in public services reflect the underlying difficulties with measuring and aggregating outputs relative to the case for market sector activities. Finally section 4 looks at evidence on some potential factors underlying private sector productivity - competition and performance incentives - and makes comparisons with studies for the public sector. Section 5 concludes.

2 Measuring outputs and inputs

Productivity measures the efficiency with which an organisation manufactures products or delivers services by measuring the relationship between an organisation's outputs and the inputs used to produce them. Productivity comparisons can be made across different individual service providers and over time, and can be used to help understand why some providers are more productive than others, or what happens to productivity when reforms such as the introduction of competition between providers are implemented.

In an ideal world from the perspective of productivity measurement, organisations would produce a single homogeneous good using a single homogeneous input. Productivity measurement then amounts to relating the number of units of the good that are produced to the number of units of the input that are used to produce it. More broadly, measuring productivity involves relating a volume measure of output to a volume measure of inputs.

In practice organisations produce more than one good or service and use a number of inputs to produce them. These goods, services and inputs may be highly heterogeneous and may differ in quality over time and across different organisations. The volume measures of outputs and inputs used in measuring productivity should ideally capture changes or differences in quality. For example, a simple measure of labour inputs such as a headcount of staff could be adjusted for the number of hours worked and the skills or human capital of those individuals.

If these quality-adjusted volume measures of outputs and inputs were obtainable, then, depending on precise method of productivity measurement used, it might be desirable to aggregate some or all of the outputs and inputs to produce measures at the level of an organisation or sector of the economy. Where inputs are shared across a range of outputs within an organisation, measuring the productivity with which a single output is produced may be problematic, for example due to difficulties with attributing the appropriate fraction of an employee's time to the delivery of one service. Comparing productivity at the provider level may mask important differences in the efficiency with which providers deliver individual outputs, but on the other hand may help detect whether overall productivity is suffering at the expense of an increase in the efficiency with which a single service, perhaps one that is being targeted, is being delivered.

Equally, productivity measurement at the sector level, such as the provision of education or healthcare services, can provide useful information on trends in productivity over time, but will mask interesting features of what is driving any productivity growth within sectors. Hence examining productivity at both the sector, and organisation level can be valuable. Research on private sector productivity growth at the sector level has examined the contributions of new entry, of expansion by more efficient providers, and of the exit of relatively poor performers in driving productivity growth, (for example Foster et al., 1998 and Disney et al., 2003).[1] When considering public sector services, exit may have significant implications for coverage and ease of access to providers, a feature of provision that is likely to be valued by society.[2] Where appropriate it is therefore important that such features are captured in output measures for public service providers so that the effects of entry and exit and expansion and contraction in provision on productivity are accurately measured.

The remainder of this section discusses the measurement of output for both private and public sectors, focusing on the problems that arise in the context of public sector services. It then covers the measurement of inputs.

[1] See Petrin and Levinsohn (2006) for a discussion of the methods used in these types of studies with regard to welfare comparisons.

[2] See Damiani et al. (2005) for an analysis based on travel times of the extent to which individuals are likely to be able to exercise choice across hospitals in England for elective care.

2.1 Measuring outputs

Both public and private sector organisations produce multiple, heterogeneous outputs. To measure the output of a private sector organisation the type of information that is typically available to researchers is a measure of the aggregate *value* of goods or services sold, often measured by gross output, sales or value-added (gross output minus intermediate inputs such as components, or goods purchased for re-sale). This aggregation across goods and services to the value of output at the organisation level implicitly involves using information on the relative prices of the goods and services as weights. In well functioning, competitive markets these prices provide information on the marginal benefit to consumers and the marginal cost of producers associated with a unit of each good or service, and hence the relative valuations of different products.

When measuring productivity these value measures of output are then transformed into volume measures using price indices. For example in making productivity comparisons over time the measure of output used should capture changes in the volume of output produced including improvements in quality, but not any changes due to inflation in the price of a unit of the good. The price of each good or service incorporates information on quality, hence price indices should ideally make adjustments for price increases due to quality improvements as opposed to inflation. In practice however this is not always feasible, and in the absence of information on prices at the level of individual products or organisations, measures of the value of output are often turned into volume output measures using industry-level price indices.[3]

While obtaining suitably dis-aggregate, quality-adjusted price indices presents problems for accurate productivity measurement in private sector organisations, for public sector organisations the problem is more severe in that there is no price information available at all. As discussed below, this raises difficulties for aggregation and measuring quality change. But it is first worth pointing out that a lack of price information is not unique to public sector services. Similar issues arise in measuring the output of financial services firms such as banks. For example in the UK customers do not always pay an explicit price for a bank account, the price being implicit though a lower interest rate. Thus the types of approaches

[3] See Foster et al. (2005) for a study that does use producer-level prices.

that have been taken to measuring the output of banks,[4] constructing count measures of numbers of accounts and transactions, have similarities with the methods used to measure the output of public sector organisations.

Measuring the output of public sector services

In deciding what it is desirable to capture in measure of the output of a public service a starting point is to consider the full range of outputs that society values and wants those services to deliver. Consider a GP's surgery. In addition to the amount and quality of the treatment provided other desirable outputs might include ease of access to treatment, such as how straightforward it is to register with a GP and the length of time patients have to wait to get an appointment. More generally public services might have equity objectives, for example that teaching improvements benefit pupils of all types, rather than just aiming to raise the average benefit.

It is also useful to distinguish between public services that at the point of use are provided to individuals, such as education and healthcare, and those that are provided collectively such as defence. Many public sector services in fact provide a mix of both, for example the police service prevents crime in addition to investigating specific criminal incidents, and both of these activities are valued. Moreover, just as individuals are willing to purchase insurance even though they may never make a claim, they may place a value on a service such as a hospital being available should they need to use it.[5]

The wide range of outputs each public service provides raises a number of questions about what information it would be desirable to collect and about how best to use the type of information typically available to construct measures of output. It is common to split potential measures into measures of activities, outputs and outcomes. In the provision of education services, these might be the number of lessons provided, the number of GSCEs obtained and the effect of those education services on the future employment and earnings of those taught respectively. Similarly, in healthcare these might correspond to the number of consultations, the number of patients successfully treated, and the improvement in health of those treated. Each offers information that can be used to measure output.

---

[4] See, for example, Fixler and Zieschang (1999).

[5] Hence unused capacity in some public services may still be considered as a valuable output.

Counts of activities are likely to be relatively easy to measure, but may not be detailed and comprehensive enough to capture all of the outputs that a service provides. Measuring the number of fires successfully extinguished by a fire service will be more straightforward than measuring fire prevention, but a good measure of the overall output of the fire service would capture both. In this case only using measurable activities to reflect overall output might provide misleading information about changes over time in the level of output, or about differences in output across providers. If a fire service in one area were more proficient at fire prevention they would consequently need to tackle a lower number of fires, and if only the latter activity were used to measure overall output, then that fire service would be incorrectly assigned a relatively low level of output. Similarly if technological developments in healthcare meant that fewer treatments were needed to deal with the same ailment, then a simple count of treatments might lead to measured output decreasing over time. One way to overcome this latter problem would be to use patients rather than activities as the unit of analysis. This output-based approach however requires information to be collected on patients, or the linking of activities to patients, rather than just collecting counts of activities carried out. See the discussion in Dawson et al. (2005) and papers cited therein.

Simple measures of activities are also unlikely to capture differences or changes in the quality of service provided. Counts of the number of lessons taught will not measure the quality of those lessons, which might be reflected better in an output measure such as the number of qualifications gained by students. In turn, the quality or value of those qualifications together with other benefits derived from education will be reflected in future outcomes for the pupils in the form of earnings and in other societal outcomes affected by education. Incorporating information on outcomes such as levels of health or crime could help to measure the quality of service provision. But this can itself pose difficult measurement problems in terms of isolating the marginal improvement in health that is due to public healthcare provision as opposed to other factors such as changes in individuals' diets that are not driven by the service provider.

Even if volumes of individual outputs of public service providers can be accurately measured, a lack of prices remains problematic from the point of view of productivity measurement. As described above for the private sector, prices, (assumed to reflect marginal benefits and marginal costs), can be used to aggregate outputs. As discussed in section 3 many of the techniques used to measure productivity require the aggregation of outputs. A key issue is

then how to weight different outputs together. In principle the weights used should reflect the marginal social benefits associated with each individual output.

There are various options for devising these weights. One often implemented in practice due to data availability is to use information on the relative costs of different outputs to aggregate. This requires an accurate attribution of costs to particular outputs which could present difficulties, for example when labour inputs are shared across a number of activities. Using cost weights means that more expensive activities receive a higher weight in the construction of an aggregate output measure, and that a change in the mix of activities carried out which moved towards more cost-effective treatments could in principle reduce measured output. The use of cost weights will only be appropriate to the extent that relative costs accurately proxy the relative marginal social benefits associated with each output.

Other possibilities that have parallels with research on measuring the value of public goods such as the environment include backing out information on relative valuations from surveys, or using information from implicit valuations, such as the premium that individuals are willing to pay to live in the catchment areas of particular schools, or how long individuals are willing to wait, or how far they are willing to travel to access a certain hospital. A further option might be to obtain price information from the private sector. However the services provided by a private sector alternative might differ in their scope and in their characteristics - for example private sector healthcare potentially offering reduced waiting times and different quality accommodation. Moreover the characteristics of those individuals using private sector alternatives, for example their underlying health, may differ from those using public sector provision.[6] Hence in each case, there would remain questions about how reliably these methods would capture the relative valuations of different goods.

A final point is whether the weight attached to a given component of output should change if society's preferences, or valuation of a unit of output change over time. Suppose that the objective is to measure productivity growth for an organisation, to determine whether or not there has been an increase in the efficiency with which it delivers goods and services. In line with productivity measurement in the private sector an increase in price of that organisation's output from one period to the next should only be reflected as an increase in the measured

[6] See also Dawson et al. (2005) for a discussion of using prices from other countries to value NHS output.

volume of output to the extent that the price increase underlies an improvement in quality. Similarly, when measuring productivity growth for a public sector organisation, any change in the marginal social valuation of its output should only translate into an increase in the real volume of output to the extent that it reflects a quality improvement, rather than any change in valuation driven by factors other than the actions of the service provider. However if the objective were to measure the nominal value of output, then changes in preferences and the relative valuations of the different goods and services produced by the organisation could change the weights used in aggregation.

Measuring public sector output in practice

Table 1 below provides examples of the types of measures and weights used in practice to measure output at the sector level. The table is derived from a recent review of the measurement of UK public sector output and productivity, 'The Atkinson Review', (Atkinson, 2005),[7] and describes measures in use at that time. The table also details some of the recommendations from the report to improve the way that the UK Office for National Statistics (ONS) measures the output of public sector services, for example to account for quality change.

From 1960s to 1998 the ONS used a convention that the value of outputs of public sector services was equal to a measure of the value of inputs, which meant that measured productivity growth was not possible. In 1998 the ONS began to construct new direct measures of output, such as those shown in the first column of table 1. For example at the time of the Atkinson Review the output of the education services sector was measured by numbers of full-time equivalent pupils enrolled at four types of maintained schools (nursery, primary, secondary and special schools) aggregated together using cost weights reflecting total UK expenditure on each type of school. A quality adjustment of +0.25 was then added annually to the cost-weighted pupil numbers for primary and secondary schools to reflect improvements in examination results.[8]

[7] In addition to the Atkinson Review see Pritchard (2002) for more detail on how the output of some public sector services is measured by the ONS.

[8] This figure is based on the trend in the average points score from GCSE results over a period of four years in the mid 1990s.

Table 1: Measuring the output of public sector services, recommendations from the Atkinson Review

Function	Main components of measure	Main recommendations going forward
Health	Hospital cost-weighted activity index, Family Health Services (number of GP consultations etc.)	Better measures for primary care; Movement towards measuring whole courses of treatment; Ideas for measuring quality change.
Education	Pupil numbers – quality adjustment of 0.25 per cent to primary and secondary schools. Cost weights by school type (nursery, primary, secondary and special schools)	Measure pupil attendance not pupil numbers; School cost weights to be updated annually; Update the quality measure for schools and develop a new extended measure which might include measuring the value of education through increased earnings; New measures of output for initial teacher training and publicly funded nursery places.
Administration of Social Security	Number of benefit claims for 12 largest benefits. No allowance for collection of contributions.	Update the index for social security administration, including adjustment for accuracy and timeliness.
Administration of Justice	Number of prisoners, legal aid cases, court cases and probation cost-weighted activity index.	More detailed measures for the criminal justice system, with possible quality adjustment to reduce value accorded to overcrowded prison cells. Further development of measures of the output of the criminal justice system as a whole.
Fire	Number of fires, fire prevention and special services. Then split into further sub-categories. Weights proportional to average staff hours spent on each sub-category.	Measure output on basis of weights that reflect the cost to the community of fire, (damage to life and property).
Personal Social Services	Children and adults in care and provision of home helps. Cost-weighted index.	Wider and more detailed coverage in the measure of adult social services output; Extension of children's social services output measure; Development work on quality adjustments

Source: adapted from Table 2.1 Atkinson Review: Final report, plus later chapters.

The Atkinson Review made a number of recommendations for improvements, such as increasing the scope of the indicators collected so that they are more disaggregated and cover a wider range of the outputs of a particular service. Examples of specific recommendations are given in the final column of table 1. For the education sector recommendations included measuring pupil attendance rather than pupil numbers, to capture more accurately the actual number of pupils being taught by accounting for absence, and this has now been implemented. A further recommendation was to develop new ways of adjusting output for quality change. Following on from the Atkinson Review the ONS (ONS, 2006a) has published new measures of education outputs and education productivity accounting for quality change using information on GSCE results and progress between Key Stages of the English education system. The same article also discusses issues relating to quality adjustments based on teaching assessments and class sizes.

A further study that ran concurrently with the Atkinson Review examined the measurement of the outputs and productivity of the National Health Service, (see Dawson et al., 2005). The Department of Health measures the output of the NHS using a cost-weighted index of activities covering hospital, primary care and community health services. The report by Dawson et al. (2005) makes methodological recommendations and suggestions for data collection to enable the construction of an output index using values as weights. This would value the health outcomes associated with different activities using information on the effect of each activity on quality adjusted life years, together with placing a value on waiting times.[9] The report also suggests ways of adjusting cost-weighted output measures for quality by incorporating factors such as survival, waiting times, re-admissions and patient satisfaction, (See also ONS, 2005b for a follow-up analysis of productivity in the health sector, and Triplett, 2001).

Many empirical studies of the health care sector also examine output and productivity in the treatment of specific medical conditions, either at the national 'industry' level or by making comparisons across providers. Cutler and McClellan (2001) and various chapters in Cutler and Berndt (2001) provide a number of examples which use information on health outcomes, such as survival rates following treatment for a heart attack, together with detailed

[9] Ideally any differences in waiting times across hospitals, or over time for the same hospital, should only be captured in an output measure for use in a productivity comparison to they extent that they are determined by the actions of the provider, rather than changes in the underlying health of patients.

information on the specific medical interventions used in treatment. Further examples are given in section 4.

Before turning to measures of inputs it is worth pointing out that in the UK National Accounts the measurement of the output of collective services such as defence, environmental protection and central government administration, follows an input equals output approach, in that is it uses information on costs of provision to measure outputs. This by definition leaves no scope for measured improvements in productivity.

2.2 Measuring inputs (and value-added)

Measuring inputs into public services such as labour and purchases of intermediate goods typically presents less of a problem, and given that such inputs are typically priced no worse a problem than measuring the equivalent inputs in the private sector. In analysing productivity in the private sector data is used on labour inputs - often a headcount of workers or the number of hours of worked and information on the cost of those workers to the firm, on the value of intermediate goods purchased for processing or resale, and on the value of the capital stock, i.e. equipment and buildings used in production.[10] Again value measures of inputs are converted into volume measures using price indices at the most dis-aggregated level available. In the public sector because these inputs will be paid for, price indices can be constructed and volume measures of inputs obtained. As before these price indices should ideally take into account quality change.

An important aspect of some public services such as education is that they can be thought of as involving joint production between the individual and the service provider. First, the characteristics of the individuals using the service will make a difference to the measure of *gross* output, and these individuals can in a sense be thought of as (intermediate) inputs themselves. For example a child's ability will affect a measure of gross output such as examination results. Second, the effort expended by an individual in conjunction with the inputs of the service provider, such as the time and effort a pupil devotes to homework and their behaviour in the classroom, may also affect output. While in the first case the innate ability of an individual is not something a school can influence, in the latter case the

[10] Dealing with durable inputs such as capital is more difficult in that a measure of capital services is preferable to a measure of the capital stock. But, for example in the case of buildings and equipment, this should be no more problematic in the public than in the private sector.

behaviour and effort put in by pupils may be. The extent to which a service provider such as a school can induce effort will be a characteristic of output which society values.

Private sector services can also involve joint production, for example car repair where a garage will vary the price it charges depending on the initial condition of the car in question. When comparing measured output across two garages, prices will serve as weights reflecting the bundle of car repair services actually supplied in each case. The desired measure is the volume of car repair services provided, not simply the number of cars repaired. However if only the latter information were available, when comparing the output of the two garages it would be necessary to take into account the complexity of the jobs undertaken at each, and only compare like with like.

This is analogous to the problem faced in the measurement of output in some public services. When making productivity comparisons in some cases it will be important to take into account the characteristics of the individuals using the service. If this is not done the results could be highly misleading. For example, it would not be accurate to ascribe a decrease in productivity to a doctor who in one month treats a large number of relatively straightforward cases and in the following month deals with a smaller but much more complex set of cases. Nor would it be desirable to compare test results from two schools where pupils had very different initial literacy and numeracy skills without making some adjustment.

In the absence of price information or weights which can capture different valuations, two options are either to adjust the measure of output directly so that it is a value-added measure,[11] or to use a gross output measure and try and take account of differences in inputs. For example, by only comparing providers who are operating in similar environments, such as schools in catchment areas with pupils of similar types, or GPs in areas with resident populations with similar demographic characteristics.

Some public services are also linked, for example primary education services are an input into secondary education. In these cases it might be desirable to separate out changes in the output of secondary education that are only due to the actions of that sector rather than any changes in the outputs of the primary education sector.

[11] For a discussion of value-added performance indicators in education see Wilson (2004).

3 Techniques for measuring productivity

An organisation's productivity is typically measured relative to some benchmark, for example compared to itself during the previous month or year, or compared to another similar organisation. Similarly productivity comparisons are made at the sector and country level. Productivity measures include single factor productivity measures which take account of only one input, such as labour productivity, and multifactor productivity measures, such as total factor productivity which might relate the volume of output to the volume of labour inputs, intermediate inputs and capital used in the production of that output.

Because of the difficulties present in measuring the output of public services, some attention has been paid to using productivity measurement techniques that do not require the aggregation of individual outputs, such as data envelopment analysis. In this section I discuss different approaches to the measurement of productivity, focussing on some applications for public sector services.[12]

3.1 Index number approaches

One commonly used method of productivity measurement is the index number approach. Below is a simple example of an index number method of measuring productivity growth, which relates a volume index of outputs to a volume index of inputs. Consider a firm producing multiple outputs y_i using multiple inputs x_i. The price of each output is p_i and the price of each input w_i. Each quantity and price is observed in two periods t and $t+1$. One way of defining output and input volume indices is as follows, where Q_o is a Laspeyres output quantity index and Q_I a Laspeyres input quantity index.[13]

$$Q_O(p^t, p^{t+1}, y^t, y^{t+1}) = \sum_{i=1}^{M} p_i^t y_i^{t+1} \Big/ \sum_{i=1}^{M} p_i^t y_i^t \qquad (1a)$$

$$Q_I(w^t, w^{t+1}, x^t, x^{t+1}) = \sum_{i=1}^{N} w_i^t x_i^{t+1} \Big/ \sum_{i=1}^{N} w_i^t x_i^t \qquad (1b)$$

[12] See also Smith and Street (2005) for a discussion of efficiency measurement in the context of public services.

[13] The Laspeyres index uses base period prices but there are many alternative approaches for example the Paasche index which uses current period prices as weights and the Fisher ideal quantity index which is the square root of the sum of the Laspeyres and Paasche indices. See Diewert (1992) for a full discussion. Index numbers can be fixed base or chain based, where in the latter case the base is updated each period.

An index measure of productivity growth is then given by the ratio of these two indices

$$\hat{a}^{t,t+1} = Q_O / Q_I \qquad (2)$$

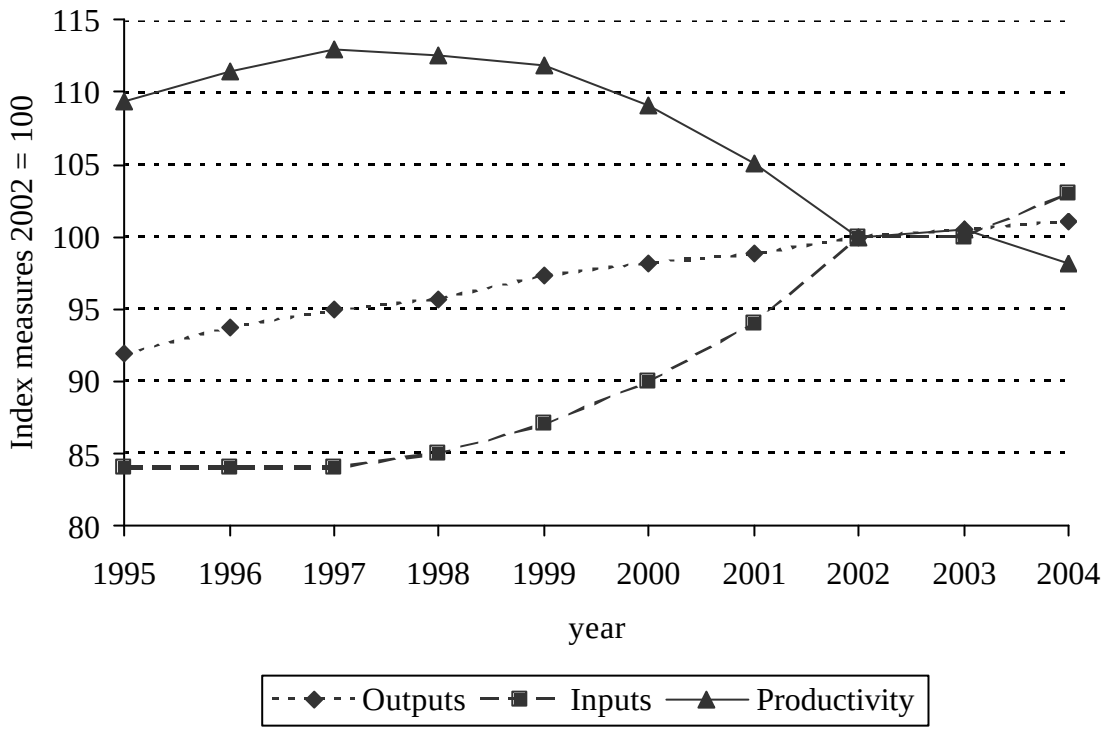
If instead of data on quantities and prices, data on revenues, input costs and prices are available then productivity indices can be constructed by deflating the revenues and costs by the relevant price indices. These are the type of data that are typically available for private sector productivity measurement.

Index measures are used in public sector productivity measurement for example at the sector level. Following the discussion in section 2.1 the missing piece of information that makes this difficult is what to use to weight different components of output together in place of the prices p_i in (1a) above. Figure 1 gives an example of index measures of output, inputs, and productivity (the ratio of the output and input indices, as in (2)) for the UK education sector. For schools output is measured using pupil attendance incorporating the +0.25 per annum quality adjustment. For nursery schools and classes output is measured using full-time-equivalent pupil numbers, and for nursery free places by the number of free places filled. Numbers of students are used to measure the output of initial teacher training courses and health professional courses. These volume measures of output are weighted together by costs.[14]

The volume of inputs is measured by deflating nominal government expenditure on education. Education expenditure includes expenditure on labour inputs (for example, teaching and support staff), expenditure on goods and services (teaching aids, electricity, transport etc.) and capital consumption (an estimate of the amount of capital services delivered in each year from durable inputs such as computers and buildings). In terms of expenditure, labour inputs account for around three quarters of total education inputs. As an example of the type of deflator used to generate measures of the volume of inputs, in figure 1 local authority pay is deflated by a local authority education pay index, and central government pay is deflated by the public sector average earnings index.

[14] Output is measured as in the National Accounts 2005.

Figure 1. Index measures of education output, inputs and productivity, UK, 2002 = 100



Source: ONS (2006a)

In each case the indices are based at 100 in 2002. As can be seen from the figure, for much of the period annual increases in the input index outstrip the corresponding increase in the output index, hence the index of productivity is falling for a large part of the period, (see ONS (2006a) for a discussion of other potential adjustments to the measurement of output and inputs for this sector).

In practice the precise form of index number used depends on the data available and underlying assumptions made by the researcher. These assumptions can, but do not have to, include economic assumptions about optimising behaviour by organisations (profit maximisation and cost minimisation), about conduct in product and input markets (perfect competition in product and input markets with organisations acting as price takers), and about the form of an organisation's production function, for example whether it exhibits constant returns to scale, and the form of production technology such as Cobb-Douglas or translog, (the latter allowing for more flexibility in the substitutability of the different inputs). An alternative to making economic assumptions to determine the form of the index is to take an 'axiomatic approach', and select an index which satisfies a set of desirable properties.

In addition to measuring productivity growth, index number techniques can also be used to make *relative* productivity comparisons across organisations. A simple example of a total factor productivity index measure based on a Cobb-Douglas technology, for an organisation i at time t is:

$$A_{it} = \ln(Y_{it}) - b_{1it}\ln(L_{it}) - b_{2it}\ln(M_{it}) - b_{3it}\ln(K_{it}) \qquad (3)$$

where $b_{1it}, b_{2it}, b_{3it}$ are the respective expenditure shares of each input - labour, goods and services or intermediate inputs, and capital - in total output, and output Y_{it}, labour L_{it}, intermediate M_{it} and capital inputs K_{it} are all expressed in volume terms.[15] A_{it} can be used to compare productivity across organisations. In this case, A_{it} also provides an absolute measure of productivity, which can be thought of as the component of the volume of output that cannot be attributed to the volume of measured inputs - a measure of technical efficiency, or as 'a measure of our ignorance' in that it is a residual. Interpreting this unexplained component of output as productivity requires assumptions of constant returns to scale, profit maximising competitive behaviour in product markets and price taking behaviour in input markets. These assumptions mean that the total value of output is equal to the total value of inputs, and that the expenditure (or cost) share of each input in total output measures its marginal productivity and forms the appropriate weight for each input. It follows that such index measures are reliant on the accurate measurement of outputs, inputs and the shares of each factor in output.

3.2 Parametric estimation

An alternative to the use of index number techniques is parametric estimation, for example estimating a production or cost function. Below is a simple example of production function estimation, which could form an alternative to the index number method given in (3). Here, instead of directly using information on the share of each factor in production the factor share of each input (now assumed to be invariant across organisations and time) is estimated.

[15] See Caves et al. (1982a,b) for discussion of superlative index numbers that allow for more flexible functional forms.

Log-linearising a Cobb-Douglas production function of the form, $Y_{it} = A_{it} L_{it}^{b_1} M_{it}^{b_2} K_{it}^{b_3}$ where A_{it} is a Hicks neutral productivity shift parameter leads to a Cobb-Douglas production function of the form,

$$\ln y_{it} = \ln A + b_1 \ln l_{it} + b_2 \ln m_{it} + b_3 \ln k_{it} + e_{it} \qquad (4)$$

which can be estimated for example using data on a number of organisations, *i*, over time, *t*. Relative productivity is obtained from the residuals e_{it}. To interpret the entire residual as a measure of total factor productivity requires assumptions of competitive product and factor markets.[16]

Productivity can also be estimated by estimating a cost function, where costs are expressed as a function of different outputs and the prices of each of the inputs. By assuming cost minimisation a cost function can be estimated together with the input factor demand equations.

As discussed further in section 4 below, in productivity analyses using data for the private sector, production functions such as that in (4) are often 'augmented' with additional variables to try and explain some component of the residual total factor productivity measure. For example data on firms' research and development (R&D) stocks might be included as an additional explanatory factor. A positive and significant coefficient on this variable would be interpreted as more R&D being associated with higher productivity. Similar types of analysis have been undertaken in the case of some public services. Krueger (1999) estimates an 'education production function', relating students' test scores (as a measure of education output) to an indicator of class size plus other student, teacher and school characteristics (which can be thought of as inputs). His results suggest a positive relationship between smaller class sizes and test scores.

[16] Production function estimation can allow the assumption of constant returns to scale to be relaxed. Both index number measures and production function estimation methods can be adjusted to allow for imperfect competition in output markets. A number of issues arise in the estimation un-biased production function parameters, for example, dealing with the endogeneity of factor inputs (firms choose their outputs and inputs simultaneously, and unobserved shocks may be correlated with both an organisation's output and input choices). Estimation issues are not discussed further here, but see for example Griliches and Mairesse (1995) and Bond and Söderbom (2005) for discussions of identification of production function parameters.

A parametric estimation technique that has been used in public sector productivity measurement is stochastic frontier analysis (SFA). Compared to the production function and cost function techniques just described, stochastic frontier analysis is concerned with modelling the efficiency, or 'production possibility frontier', rather than capturing the mean of the distribution. SFA essentially splits the error term in equation (4) into two components, or two types of deviation from the production possibility frontier, one is a measurement error or noise component h_{it}, and the other is an inefficiency component u_{it} [17]

$$\ln y_{it} = \ln A + b_1 \ln l_{it} + b_2 \ln m_{it} + b_3 \ln k_{it} + (h_{it} + u_{it}) \qquad (5)$$

On average the noise component is zero and the inefficiency component is expressed as a deviation from the estimated efficient frontier. This technique requires assumptions about the joint distributions of the two components and the functional form of the frontier. The data are uninformative about these issues, but the resulting efficiency estimates will be sensitive to the assumptions made.

Stochastic frontier analysis has been used in a 'cost frontier' form to estimate relative efficiency in a number of applications for the public sector. This relates an organisation's costs to input prices and the volumes of outputs it produces, and measures the inefficiency of each organisation relative to the estimated least cost frontier. This approach has attractions in public sector productivity measurement because it does not require the aggregation of each organisation's outputs.

Examples of applications of SFA techniques to measure the relative efficiency of English hospitals include Jacobs (2001) and Street (2003). These papers compare efficiency rankings generated using SFA techniques to those obtained using alternative methods such as cost function estimation. Street (2003) relates a measure of total costs incurred by a hospital to a number of hospital characteristics capturing the activities carried out and the characteristics of patients, for example the number of first accident and emergency attendances, the number of first outpatient attendances, the number of inpatients adjusted for the complexity of the case-mix, and the proportion of patients aged 60 years and over. He finds that while a high overall correlation is achieved between the efficiency rankings assigned by the different

[17] The production function estimation technique in (4) measures productivity, (as given by the error term which is assumed to follow a normal distribution), relative to a mean value of zero.

estimation methods, this masked cases where individual hospitals were given substantially different rankings according to the technique used, and the assumptions made under different SFA specifications. This leads to the conclusion that it would be unwise to base any hospital efficiency targets on rankings derived from a single efficiency estimation technique. This echoes the concerns of Newhouse (1994), who concludes that frontier estimation techniques such as SFA, combined with imperfect data on inputs and outputs, should not be used to determine re-imbursement rates for hospitals.[18]

3.3 A non-parametric technique

A further technique that has been proposed for measuring the efficiency of public sector organisations is data envelopment analysis (DEA).[19] DEA is a non-statistical approach, which again tries to model the efficiency or production possibility frontier. It takes data on organisations' outputs and inputs, and measures the efficiency of a particular organisation by its distance from the 'outer envelope' of the data. This outer envelope is assumed to measure the combination of outputs that a fully efficient organisation could deliver for a given set of inputs, and all deviations from the frontier are classed as inefficiency.

The technique relies on the use of extreme observations in the data, (e.g. the maximum amount of a particular output), hence in determining the position of the frontier and the individual efficiency scores it is very sensitive to mis-measurement in the data. A further problem for the accuracy of the results is that if the data are sparse in places, and no real comparator organisations are observed, it is possible to end up comparing organisations to only hypothetical efficient organisations on the frontier.

The attraction of this technique for the measurement of public sector productivity is that it does not require information on weights to aggregate outputs (or inputs), and does not require information on prices. It essentially allows the data to determine the weights, (subject to any restrictions placed on them by the researcher) so that an organisation's productivity is presented in the best possible light.

---

[18] See also World Health Organisation (2000), which used SFA techniques to make comparisons of the relative performance of different countries' healthcare systems, and the subsequent critical assessment of the study.

[19] See for example Spottiswoode (2000), which proposed the use of data envelopment analysis, and stochastic frontier analysis for the measurement of police force efficiency. See also the discussion in Stone (2002) outlining why use of these techniques may be misguided.

The approach still requires some assumptions to be made, such as whether production is subject to constant or variable returns to scale. Under variable returns to scale the technique can designate an organisation as fully efficient simply because it produces more of a particular output than other organisations. Hence the results will be sensitive to extent to which outputs and inputs are aggregated, with more disaggregated outputs likely to lead to more organisations being classed as 100% efficient. This may create a trade off, given that the attractive feature of the technique was that the ex-ante aggregation of outputs was not required. A final but important issue is that the weights used in DEA may not reflect in any way the relative marginal social valuations of different outputs.

Examples of the use of DEA in practice include Ruggiero (1996) in the context of measuring the efficiency of education provision, Jacobs (2001) looking at hospital efficiency and Førsund et al. (2006) who examine the efficiency of tax offices. Ruggiero (1996) constructs measures of education production efficiency for New York school districts. Outcomes, comprising test scores for reading, mathematics and social studies, are related to inputs including measures of teacher aides and assistants per pupil, a measure of teacher quality, computers and classrooms per pupil, plus an environmental factor based on the poverty rate in the district, which is assumed to be unaffected by the behaviour of education providers.[20] The paper highlights the sensitivity of the efficiency scores produced to different methods used to generate them. Førsund et al. (2006) in their investigation of the efficiency of tax offices also stress the importance of subjecting the results of DEA to a sensitivity analysis. Jacobs (2001) does not find strong correlations between efficiency scores for hospitals generated by SFA and DEA techniques. All these studies therefore highlight the importance of ascertaining robustness for efficiency measures derived using this type of technique.

3.4 Partial efficiency measures and performance indicators

It is also possible to use partial efficiency measures to capture organisations' productivity in terms of delivering individual outputs or services. Measuring productivity for a single output clearly gets around the problem of specifying weights to aggregate outputs together, but may in turn create difficulties in terms of isolating the precise amount of inputs used to deliver that particular output, for example the number of hours a nurse devotes to one specific activity.

[20] This therefore assumes that there is no sorting of individuals into school districts according to school performance.

Hence there may potentially be a trade off between using an accurate measure of a particular output or outcome and a precise measure of inputs. However studies such as those examining output and productivity growth for individual medical conditions such as heart attacks can incorporate detailed information on the specific treatments used and the characteristics of individuals receiving treatment.

Partial efficiency measures may be relatively easy to interpret and monitor which suggests that they may provide useful indicators of performance. But because such measures are only partial they may have drawbacks if used in regulation or as a tool in performance evaluation. They will not capture all of the outputs delivered by an organisation, and constructing an overall evaluation of an organisation's productivity may prove difficult if it scores highly on some measures but poorly on others. Moreover there are concerns that if particular indicators are used to assess providers' performance perverse incentives may be created, with measured outcomes being focussed on at the expense of performance in other un-measured activities, and potential incentives for organisations to select those individuals to provide services to who will best enable them to meet performance indicators, for example patients who are relatively straightforward to treat.[21]

4 What underlies productivity growth and differences in productivity across organisations?

Studies that only measure productivity growth or make efficiency comparisons across organisations sometimes shed little light on what drives changes or differences in productivity. This section briefly discusses some empirical research that has sought to identify factors that affect productivity and productivity growth in both the private and public sectors. Two factors that have received attention in research on both sectors are competition and performance incentives for employees.

4.1 Competition

The empirical evidence on the relationship between competition and productivity differs somewhat across the private and public sectors. Given the differences between the incentives

[21] For discussions of the use of performance measures in the public sector see for example Propper and Wilson (2003), and Stevens et al. (2006).

and constraints facing firms in private markets and those affecting the operation of public sector organisations the fact that the studies may come to different conclusions is not necessarily surprising. However because of the inherent differences between the sectors and the difficulties in measuring productivity for public sector organisations, empirical studies for the two sectors tend to use quite different data and measures, making comparability less straightforward. A key issue for research in this area is the difficulty of picking up the causal relationship between competition and productivity. To do this some studies of the effects of competition in both the private and public sectors exploit policy reforms that affect the degree of competition organisations face.

In private markets if individuals and firms can choose between alternative providers of goods and services competition provides suppliers with incentives to increase efficiency, for example if they are operating with a degree of slack, or else risk going out of business. Competition may also induce higher productivity growth or improve dynamic efficiency by providing firms with greater incentives to innovate, such as by creating new products and services to try and reduce the amount of competition they face from other firms in the market.

Empirical studies of the relationship between private sector productivity and the extent of competition in markets for goods and services tend to find that competition improves productivity growth. These studies typically relate a measure of productivity to measures of the extent of product market competition in a regression framework, or add in measures of product market competition as extra variables in a production function such as in (4) above. Nickell (1996) using firm level data examines the relationship between both the level of total factor productivity and growth in total factor productivity and a range of measures of market structure and product market competition, (e.g. measures of market share, the level of rents earned by firms and the degree of import penetration). He finds evidence that more intensive product market competition is associated with higher productivity growth. Disney et al. (2003) carry out a similar analysis and find evidence that greater competition is associated with both a higher level and growth rate of productivity. See Aghion and Griffith (2005) for a survey of theoretical and empirical work in this area.[22]

[22] See also the literature on the efficiency effects of privatisation and deregulation for example Green and Haskel (2004).

Analyses of productivity growth in private sector markets have also pointed to the role of entry, exit, expansion and contraction of firms in productivity growth. Disney et al. (2003) highlight the role of restructuring. Using data on UK manufacturing establishments they show that plants that exit had on average lower productivity than surviving establishments and that entrants exhibited higher productivity levels. See also Pavcnik (2002). This type of evidence points to a significant role of restructuring and the reallocation of resources towards more efficient providers in driving aggregate private sector productivity improvements.

For public sector organisations flexibility in terms of the scope for expansion and contraction of activities, and the incentives to adjust, may be more limited. While the closure of a relatively poorly performing firm in private sector markets may raise the average level of productivity in the industry, this might be less clear cut in the case of public sector organisations if an important component of output is accessibility. Such circumstances re-enforce the importance of capturing all of the elements of output that are valued by society in a productivity measure. If access to services were not part of the measure of the volume of output then it may potentially be incorrectly inferred that productivity had risen, when in fact it had declined.

For public sector services the evidence on the effects of competition on efficiency is more mixed than for the private sector.[23] Due to the difficulties of measuring productivity for public sector organisations some studies have focused either on partial output measures, (providing a good measure for a single outcome rather than a measure of the total output of an organisation), or on partial measures of cost-effectiveness. One such study for the healthcare sector is Kessler and McClellan (2000). This study looked at the relationship between competition and two outcomes - mortality rates post admission for heart attack patients and the healthcare costs associated with treatment, using data from the US. The measure of competition used is based on travel distances between patients and hospitals, and captures the intensity of competition or choice within geographic markets. For the period post-1990 they find evidence that competition resulted in lower death rates and lower costs, while their evidence for the pre-1990 period studied suggested that greater competition was associated with higher costs. Propper et al. (2004) look at the effects of competition on an

[23] See Burgess et al. (2005) for a more detailed assessment of the economic evidence on the effects of choice and competition in education and healthcare.

outcome measure. They find that during the NHS internal market, (which was associated with a greater degree of competition between hospitals), competition was associated with a decrease in the quality of outcomes as measured by higher death rates following hospital admission for a heart attack. Even if this were accompanied by an increase in productivity it seems unlikely to be a desirable outcome from a welfare perspective.

Burgess et al. (2005) conclude that the empirical evidence suggests that competition between hospitals is associated with lower costs, while the evidence of the effects on the quality of outcomes is more mixed.[24] Given the difficulties in measuring all dimensions of output at the level of an organisation such as a hospital, and hence such a measure being potentially misleading, it is not surprising that many studies look at single outcome measures that may be more reliable. But given that such studies do not assess the effects of competition at the level of an organisation as a whole, together with the absence of data on inputs in some studies, it is difficult to make direct comparisons with the empirical evidence for the private sector.

Burgess et al. (2005) also survey the empirical literature on the effects of competition on schools. There is some evidence for the US that the threat of losing pupils induced by greater competition is associated with an increase in school productivity, as measured for example in Hoxby (2003) by an indicator of student achievement for a particular grade relative to real expenditure per student, (although see also the debate and mixed results in Hoxby, 2000 and Rothstein, 2005). Clark (2006) exploits a reform to UK schools to investigate the effect of greater school autonomy on performance. His study also looks at effects of increased competitive pressure from the newly autonomous schools on their neighbours, and finds little evidence that neighbouring schools improved their performance significantly as a result.

4.2 Performance incentives

Many studies of the relationship between the use of performance based remuneration and firm productivity in the private sector show a positive association between the two. Studies that relate the use of performance related pay to organisation level productivity often look at the use of (sometimes tax advantaged) employee or executive ownership of shares or share options, which link individuals' income to the overall performance of the organisation that they work for. For example Conyon and Freeman (2004) estimate a firm-level production

[24] See also Gaynor (2006), for a survey of the effects of competition in healthcare markets.

function, (as described in section 3.2), augmented with measures of the use of shared compensation schemes and find a positive association between some forms of employee share ownership and firm productivity. Again, what is tricky in this literature is isolating causal effects; it may simply be the case that more productive firms use performance related pay or that there is some other characteristic of such firms that is difficult to measure, such as managerial talent, that is reflected in both higher productivity and the use of such incentives. See Prendergast (1999) for an assessment of a wide range of empirical studies in this area.

Other incentives schemes used in the private sector which do not measure performance on the basis of profit or sales but instead use other performance targets may be closer comparators to the types of scheme it might be feasible to implement in the public sector. One study of such a scheme for the private sector is that by Lazear (2000). Analysing a particular company he finds that a change from an hourly wage to piece rate based pay has a substantial effect on productivity as measured by output per worker.[25] However, implementing such schemes in public sector organisations relies on being able to measure performance accurately.

Two examples of studies for public sector organisations also indicate some positive effects of performance incentives on outputs.[26] Both of these studies exploit reforms and both examine multiple output measures. Burgess et al. (2004) look at the impact of team based performance pay on output in a large UK public sector agency. They find that for small teams where monitoring will be easier and free riding less prevalent, the introduction of performance related pay was associated with an increase in the main measure of output. Atkinson et al. (2004) assess the effects of a performance related pay scheme for teachers and look at effects on both a gross output measure (test scores) and on value-added, and also examine whether effects differed across subjects. They find that the scheme did act to improve test scores and value-added. See also Marsden and Belfield (2006) for an assessment of the same scheme.

While such studies allow for a careful assessment of heterogeneous effects on different outputs, the problems of measuring output and productivity at the organisation level for public sector services means that it is difficult to gauge the full, overall response to such incentives. This will be particularly important in the case of assessing this type of reform, as

[25] See Bandiera et al. (2006) for an analysis of the effects of incentives for managers on the performance and composition of their employees using an experimental research design.

[26] See also Burgess and Ratto (2003) for a survey of evidence in this area.

concerns might arise regarding effort being diverted to those activities that are being targeted by performance measures at the expense of other tasks.

In addition for public sector services an important question to address might be whether any effects of competition or of the introduction of performance incentives on service delivery vary across different types of individuals using the services. The extent to which different individuals, for example with different ability levels or in different health, benefit equally from a public sector service might be an outcome that society values, but will not be captured well in an average outcome measure.

In summary while some, (but not all), of the evidence on the effects of competition and performance incentives points in the same direction for both private and public sector organisations, there remain differences in the data and methods applied in studies of the two sectors which may mean that the results are not always comparable. This is primarily due to the inherent difficulty of measuring all of the features of the output of public sector organisations that are valued by society. A move towards such measures would make for greater comparability with private sector studies and provide a better picture at the organisation and more aggregate levels, but this may trade off the accuracy of single outcome measures that can encompass quality. A problem for studies in both sectors is isolating causal effects on productivity. Reforms to public services, such as the introduction of competition, and the use of pilot schemes present some of the best opportunities for research to do so.

5 Conclusions

Measuring output for public sector services is problematic, both in terms of capturing all the various dimensions of output that society values, and in measuring the relative valuations of each dimension to construct aggregate measures. In the absence of accurate, all encompassing measures of output productivity measurement risks being uninformative.

Empirical studies of productivity for public sector organisations have demonstrated that efficiency measures and rankings can be sensitive to the techniques used to derive them, and that some may not be particularly robust. This is not to say that productivity measurement for public sector organisations is not worthwhile. Productivity measurement for private sector organisations also presents a number of difficulties, but what is important is that the results of

studies, such as those examining the effects of competition on productivity, are demonstrated to be robust to the use of different productivity measurement techniques.

Partial measures of productivity, for example for the treatment of specific health conditions, can be very accurate in terms of capturing output quality. But improving measurement at the level of public sector organisations as a whole will be important to understand fully the effects of competition and other factors on output, outcomes and productivity. This would also be valuable as reforms to public services and the use of pilot schemes produce good opportunities for researchers to understand these relationships better.

References

Aghion, P. and Griffith, R. (2005) *Competition and Growth: Reconciling Theory and Evidence*, MIT Press, London, U.K.

Atkinson, A. et al. (2004) 'Evaluating the Impact of Performance Related Pay for Teachers in England' CMPO working paper 04/113.

Atkinson, T. (2005) *Atkinson Review: Final report, Measurement of Government Output and Productivity for the National Accounts*, HMSO, Palgrave Macmillan. http://www.statistics.gov.uk/about/data/methodology/specific/PublicSector/Atkinson/final_report.asp

Bandiera, O., Barankay, I. and Rasul, I. (2006), 'Incentives for Managers and Inequality Among Workers: Evidence from a Firm Level Experiment', *Quarterly Journal of Economics*, forthcoming.

Bond, S. and Söderbom, M. (2005) 'Adjustment Costs and the Identification of Cobb-Douglas Production Functions' IFS working paper WP05/04.

Burgess, S., Propper, C., Ratto, M. and Tominey, E. (2004) 'Incentives in the Public Sector: Evidence from a Government Agency' CMPO working paper 04/103.

Burgess, S., Propper, C. and Wilson, D. (2005), '*Choice. Will More Choice Improve Outcomes in Education and Health Care? The Evidence from Economic Research*', http://www.bris.ac.uk/Depts/CMPO/choice.pdf

Burgess, S., and M. L. Ratto (2003) 'The Role of Incentives in the Public Sector: Issues and Evidence', *Oxford Review of Economic Policy*, 19(2).

Caves D., Christensen L., and Diewert E., (1982a) 'The Economic Theory of Index Numbers and the Measurement of Input, Output and Productivity', *Econometrica*, 50(6), pp.1393-1414.

Caves D., Christensen L., and Diewert E., (1982b) 'Multilateral Comparisons of Output, Input and Productivity Using Superlative Index Numbers', *Economic Journal*, 92, pp.73-86.

Clark, D. (2006), 'The Performance and Competitive Effects of School Autonomy' mimeo University of Florida.

Conyon, M. J. and Freeman, R.B. (2004), 'Shared Modes of Compensation and Firm Performance: U.K. Evidence' in Card, D., Blundell, R. and Freeman, R.B., (eds.) *Seeking a Premier League Economy. The Economic Effects of British Economic Reforms, 1980-2000*, Chicago, University of Chicago Press.

Cutler, D. and Berndt, E. (eds.) (2001), *Medical Care Output and Productivity*, Chicago, University of Chicago Press.

Cutler, D. and McClellan, M. (2001), 'Productivity Change in Health Care', *American Economic Review*, 91(2), pp. 281-286.

Damiani, M., Propper, C. and Dixon, J. (2005) 'Mapping Choice in the NHS using Routine Data', *British Medical Journal*, 330, 284.

Dawson, D. et al. (2005). 'Developing New Approaches to Measuring NHS Outputs and Productivity, Final Report', CHE Research Paper 6, NIESR Discussion Paper DP.264. http://www.niesr.ac.uk/pdf/nhsoutputsprod.pdf

Diewert, W.E. (1992) 'The Measurement of Productivity', *Bulletin of Economic Research*, 44(3), 163-198.

Disney, R. Haskel, J. and Heden Y. (2003) 'Restructuring and Productivity Growth in UK Manufacturing', *Economic Journal*, 113, pp.666-694.

Fixler, D. and Zieschang, K. (1999), 'The Productivity of the Banking Sector: Integrating Financial and Production Approaches to Measuring Financial Service Output', *Canadian Journal of Economics*, 32(2), pp.547-569.

Førsund et al. (2006), 'The Tax Man Cometh – But Is He Efficient?' *National Institute Economic Review*, 197, July.

Foster, L. Haltiwanger, J. and Krizan, C. (1998), 'Aggregate Productivity Growth: Lessons from Microeconomic Evidence', NBER Working Paper, 6803.

Foster, L. Haltiwanger, J. and Syverson, C. (2005), 'Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?' NBER Working Paper 11555.

Gaynor, M. (2006), 'What Do We Know About Competition and Quality in Health Care Markets?' CMPO working paper 06/151.

Green, R. and Haskel, J. (2004), 'Seeking a Premier League Economy: The Role of Privatization' in Card, D., Blundell, R. and Freeman, R.B., (eds.) *Seeking a Premier League Economy. The Economic Effects of British Economic Reforms, 1980-2000*, Chicago, University of Chicago Press.

Griliches, Z. and Mairesse, J. (1995), 'Production Functions: The Search for Identification', NBER working paper 5067.

Hoxby, C. (2000) 'Does Competition Between Public Schools Benefit Students and Taxpayers?' *American Economic Review*, 90(5), pp.1209-1238.

Hoxby, C. (2003), 'School Choice and School Productivity: Could School Choice be a Tide that Lifts all Boats?' in Hoxby, C. (ed.), *The Economics of School Choice*, Chicago, University of Chicago Press.

Jacobs, R. (2001), 'Alternative Methods to Examine Hospital efficiency: Data Envelopment Analysis and Stochastic Frontier Analysis', *Health Care Management Science*, 4, pp.103-115.

Kessler, D.P. and McClellan (2000), 'Is Hospital Competition Socially Wasteful?', *Quarterly Journal of Economics*, 115(2), pp.577-616.

Krueger, A.B. (1999) 'Experimental Estimates of Education Production Functions', *Quarterly Journal of Economics*, 114(2), pp.497-532.

Lazear, E.P. (2000), 'Performance Pay and Productivity', *American Economic Review*, 90(5), pp.1346-61.

Marsden, D. and Belfield, R. (2006). 'Pay for Performance Where Output is Hard to Measure: the Case of Performance Pay for School Teachers', CEP discussion paper no.747.

Newhouse, J. (1994) 'Frontier Estimation: How Useful a Tool for Health Economics?' *Journal of Health Economics*, 13, pp.317-322.

Nickell, S. (1996), 'Competition and Corporate Performance', *Journal of Political Economy*, 104(4), pp.724-746.

ONS (2006a), Public Service Productivity: Education, *Economic Trends*, 626, January, pp.13-37. http://www.statistics.gov.uk/articles/economic_trends/ET626_CeMGA.pdf

ONS (2006b), Public Service Productivity: Health, *Economic Trends*, 628, March, pp.26-57. http://www.statistics.gov.uk/articles/economic_trends/ET628_Productivity_Heath.pdf

Pavcnik, N. (2002), 'Trade Liberalization, Exit and Productivity Improvements: Evidence from Chilean Plants', *The Review of Economic Studies*, 69, pp. 245-76.

Petrin, A. and Levinsohn, J. (2006) 'Measuring Aggregate Productivity Growth Using Plant Level Data', University of Chicago, GSB and University of Michigan, mimeo.

Prendergast, C. (1999), 'The Provision of Incentives in Firms', *Journal of Economic Literature*, 37(1), pp.7-63.

Pritchard, A. (2002) 'Measuring Productivity Change in the Provision of Public Services', *Economic Trends*, 582, May, pp.20-32.

Propper, C., Burgess, S. and Green, K. (2004), 'Does competition between hospitals improve the quality of care? Hospital death rates and the NHS internal market' *Journal of Public Economics*, 88(7-8), pp.1247-1272.

Propper, C. and Wilson, D. (2003), 'The Use and Usefulness of Performance Measures in the Public Sector', *Oxford Review of Economic Policy*, 19(2), pp.250-267.

Rothstein, J. (2005). 'Does Competition Among Public Schools Benefit Students and Taxpayers? A Comment on Hoxby (2000)', *American Economic Review*, forthcoming.

Ruggiero, J. (1996) 'Efficiency of Educational Production: An Analysis of New York School Districts', *Review of Economics and Statistics*, 78(3), pp.499-509.

Smith, P.C. and Street, A. (2005), 'Measuring the Efficiency of Public Services: the Limits of Analysis', *Journal of the Royal Statistical Society Series A*, 168(2), pp.401-417.

Spottiswoode, C. (2000), *Improving Police Performance: A New Approach to Measuring Police Efficiency*, HM Treasury, London, U.K.

Stevens, P., Stokes, L. and O'Mahony, M. (2006), 'Metrics, Targets and Performance', *National Institute Economic Review*, 197, July.

Stone, M. (2002) 'Can Public Service Efficiency Measurement be a Useful Tool of Government? The Lesson from the Spottiswoode Report', *Public Money & Management*, July-September, pp.33-39.

Street, A. (2003) 'How Much Confidence Should We Place In Efficiency Estimates?' *Health Economics*, 12, pp.895-907.

Triplett, J. (2001), 'What's Different about Health? Human Repair and Car Repair in National Accounts and in National Health Accounts' in Cutler, D. and Berndt, E. (eds.), *Medical Care Output and Productivity*, Chicago, University of Chicago Press.

Wilson, D (2004) 'Which Ranking? The Impact of a 'Value-Added' Measure of Secondary School Performance' *Public Money and Management*, 24(1), pp.37-45.

World Health Organisation (2000), *The World Health Report 2000. Health Systems: Improving Performance*, World Health Organisation, Geneva, Switzerland. http://www.who.int/whr/2000/en/whr00_en.pdf, and accompanying 'World Health Report 2000 Consultation and Peer Review' http://www.who.int/health-systems-performance/consultation.htm