Does national policy have an impact on NHS productivity?

Bojke C, Castelli A, Grasic K and A Street

Centre for Health Economics, University of York

Paper to be presented at the HESG meeting (A059), Sheffield, 8-10 January 2014

Abstract

Aims

Over the last 14 years the NHS has been subject to numerous policy changes: substantial funding increases, austerity, successive re-organisations, reform of payment arrangements between commissioners and providers, waiting time initiatives, revisions to staff terms and conditions, and investments in the capital stock. In this paper we consider the impact of these policy changes on the overall productivity of the English NHS.

Methods

Productivity change is measured by comparing year-on-year growth in output against growth in inputs. Output comprises the total volume of services provided to all NHS patients treated in all settings. Quality of care is measured by waiting times, survival rates, health outcomes and blood pressure control. Inputs include the staff, equipment and capital resources that contribute to the production of care.

Data

Data come from the Hospital Episode Statistics, Reference Costs returns, Patient and Household Surveys, Prescription Pricing Authority, Electronic staff record and NHS organisational accounts.

Results

Between 1998/99 and 2011/12 the volume of NHS activity has increased by 79%. Survival rates after hospitalisation increased from 97.2% to 98.2% and blood pressure management has improved in primary care. Waiting times reached their lowest levels in 2008/9, but have increased subsequently. On the input side there have been increases in staffing levels, though these have flattened out recently, and large increases in the use of intermediate and capital inputs. Overall, input growth increased by 78% between 1998/99 and 2011/12.

While year-on-year productivity growth rates are sometimes positive and sometimes negative, looking over a succession of years reveals that output growth and input growth track each closely. Over the full 14 years, output growth and input growth have tracked each other closely and negative productivity growth has since been offset by periods of constant and positive productivity growth.

Conclusions

The NHS is in a constant state of flux and it is difficult to isolate the impact of any particular policy. The data suggest that productivity has remained largely unchanged with shifts in inputs leading to corresponding broadly constant returns to scale shifts in outputs (possibly with a lag). The NHS appears to respond to direct initiatives such as waiting time targets, but more indirect theory-based changes (e.g. provider/commissioner split) appear to have no observable impact on productivity.

1. Introduction

Policy makers invariably claim that their reforms will improve the performance of the health care system. This might be achieved by targeted measures (i.e. waiting times, contractual reforms) or by more far-reaching changes (eg general funding increases). Sometimes reforms are introduced simultaneously, often they run concurrently, frequently they will be inter-related. It is not straightforward to identify the effects of each reform, particularly when they produce spillover or unintended consequences, nor to disentangle the separate effect of each reform. We are not going to attempt to do this. Instead, in this paper we construct a summary measure of the overall productivity of the healthcare system over the last 14 years, and consider changes in productivity in the context of the various reforms taking place over the same period.

This paper is organised as follows. Section 2 briefly presents a number of health care reforms which took place in the NHS in England between 1997/8 and 2011/2 and outlines the effects that these policies may have had on (a) total volume of activity, (b) the quality of care of NHS output, (c) the amount of NHS inputs used and (d) input prices. In section 3, we first set out the theoretical framework for the measurement of productivity; we then present the specific indices of NHS output, input and productivity growth. Data sources used to populate both the NHS output and input series are described in Section 4. We report and discuss the NHS output, input and productivity growth series in Section 5. This paper concludes with an open discussion section and questions for future research.

2. Policy changes in England

Over the last 14 years the NHS has been subject to numerous policy changes. These have ranged from substantial increases in the funding of NHS activity, austerity, successive re-organisations, reform of payment arrangements between commissioners and providers, waiting time initiatives, revisions to staff terms and conditions, and investments in the capital stock. In this paper we consider whether these policy changes may have had an impact at all on the overall productivity of the English NHS.

<Refer to page 14>

Page 14 shows the timeline for the different policy changes and NHS re-organisations that have taken place in the NHS from 1997 onwards. In the following section we summarise a number of the major policy initiatives and detail the intended (as per policy documents) effects that these policies should exert on productivity and efficiency and whether it is possible to determine the source of efficiency gains via a predicted impact on either NHS inputs, NHS outputs or both.

2.1 Policy changes affecting outputs

Reforming NHS Financial Flows (DH 2002) introduced the new financial framework of Payment by Results (PbR) – a process in which hospitals were to be paid for the activity actually undertaken rather than via a block contract. Under PbR, hospitals are paid a fixed national tariff for each patient treated and they face no explicit controls on the volume of activity they perform. Hospitals have a financial incentive to treat ever more patients as long as they believe that the tariff exceeds the cost of treating an additional patient (Mannion, Marini et al. 2008). It was hoped that Primary Care Trusts would be able to resist the ambition of hospitals to undertake ever more elective activity.

Initially PbR covered only a narrow range of hospital activity but was progressively rolled out to cover the majority hospital activity. Some hospitals, though, could elect to have all of their activity paid for under PbR. Such hospitals were Foundation Trusts (FTs), the first of which were established in 2004 (*Delivering the NHS Plan,NHS,2002)*, with progressively more being afforded FT status over time. As well as facing a different payment regime, FTs had greater financial autonomy than ordinary hospitals, and were allowed to create and retain surpluses through efficient provision of services. If the underlying theory is correct then this sector of the NHS should become increasingly more efficient over the period in which the number of FTs grows and this could potentially be observable in productivity growth measures.

A further desired effect of PbR was that, by fixing prices (and thus negating the possibility of price competition), hospitals faced sharper incentives to improve quality. However since quality is not always easy to observe then it may be plausible for hospitals to reduce the costs of providing activity by lowering (costly) quality.

Waiting times targets, introduced in the 2004 *(Department of Health 2004): Putting people at the heart of public services*, may also have a direct and predictable impact on output. Most obviously if the stock of patients needed to be treated is unaffected by the policy, then improving the flow leads to increased volume. Further to that, though, waiting times may be seen as a component of the quality of output and thus reducing waiting times should not only increase quantity, but also quality.

It is worth noting a fundamental difference between policies such as PbR and waiting time directives. PbR sets up an environment in which decision makers have an incentive to be efficient, whereas target based policies spell out outcomes precisely. These policies were somewhat complementary: the waiting time target would have helped focus PbR incentive to increase activity on those on the waiting list.

2.2 Policy changes affecting inputs

In tandem with the introduction of the internal market, the NHS also faced huge changes in the way that capital was accounted for. The *NHS and Community Care Act 1990* made capital financing the responsibility of hospitals trusts. All existing capital assets were transferred to hospital trusts as liabilities with the government as creditor and only shareholder. Trusts were required to make annual surpluses of revenue over expenditure of about 6% of the value of their assets and to make a charge for depreciation. By making trusts pay directly for the use of such capital the overall objective was to give them greater incentives to manage capital more efficiently. This might have involved selling off capital considered either surplus to requirements or with too high an alternative commercial value, for example city centre land (Talbot-Smith and Pollock 2006). This shift in policy ultimately led to the Private Finance Initiative (PFI) under which hospital trusts commission private sector enterprises to build, maintain and operate all non-clinical services for a long period of time, with specific clauses set out in a legal contract. In return, the trust pays an annual fee, which covers the cost of building and maintaining the hospital, and the cost of the additional non-clinical services, such as laundry and catering, provided by the private firm *(Sahoul, Stafford et al. 2011).*

From 1997 to 2004, PFI accounted for 98% of building schemes in the NHS hospital and community health services (Talbot-Smith and Pollock 2006).

There have been major changes to labour contracts for GPs and hospital consultants with the new *General Medical Services contract* for GPs (negotiated and signed in June 2003, effective from 1st April 2004) and the new *hospital consultant contract* (2003). The new GMS contract allowed GPs more flexibility in what they did. For example GPs were allowed to opt out from 24 hour care responsibility for their patients with an income decrement. But via the introduction of the Quality and Outcomes Framework (QOF), they could generate further payments when achieving a set of targets specifically designed to improve the quality care to their patients. The immediate impacts of the new GP contract were (1) large numbers of GPs taking the option of dropping their out-of-hours commitment and (2) an increase in average GP salary of 20% per annum (http://news.bbc.co.uk/1/hi/health/4373519.stm) compared to a national average of 4.5% despite This implies that QoF quality incentives were effective in making GPs respond to quality based targets, though it is not clear that it will have led to a reduction in the gap in clinical achievement between practices operating in more affluent and less affluent areas.

The 2003 consultant contract placed a greater emphasis on planning (where programmed activities such as inpatient duties or audit or outpatient clinics, etc. are allocated to available time slots) and thus led to increased potential for managing consultants' activity. However (Bloor, Freemantle et al. 2012) report that there was no direct link to pay for performance, other than a pre-existing clinical excellence award scheme. They find that despite DoH expectations of a 1.5% productivity gain per annum, the evidence suggested that whilst hospital consultant salaries rose by 27% in the first three years following the new consultant contract and that there was no impact on a generally decreasing trend in activity.

In addition to the new contracts for GPs and consultants, Agenda for Change (2004) required all NHS staff (except doctors and dentists and the most senior executives at or just below board level) to undergo annual development reviews. These reviews were intended to compare actual individual level performance and human capital requirements against a set of key requirements for their specific post as identified using the NHS 'knowledge and skills framework'. The intention of the policy was to improve the human capital element of the workforce and thus lead to general improvements in productivity such as increased number of patients treated, reduction in waiting time and improved quality of care (Agenda for Change Project Team 2004).

3. Methods

(Triplett 2011) argues that the conventional framework of productivity and productivity growth measurement rest inherently on the concept of an underlying production function – a systematic relationship between output and the inputs used to create it. This paper takes our extensive work on the measures of productivity growth in the English NHS from 1998 to 2012 and attempts to relate these estimates back to this underlying concept. In particular we look at: the extent to which Government policy has influenced either the shape of the production function and/or movements along the function; and the ability of conventional measures of productivity growth to pick up the impact of policy.

3.1 Conceptual Model

To underpin the analysis of productivity measurement in a general setting, (Coelli and Prasada Rao 1998) define an economic-theoretic framework (aka functional approach) based on an underlying production function approach. There are conceptual and practical challenges that must be faced in applying a general productivity framework to measure health system productivity. Two are particularly salient. First, there is a multitude of diverse outputs and inputs involved in the production process. Dealing with this requires the construction of output and input indices, and we shall return to this matter shortly. Second, no *single* health production function is likely to exist. Rather it will be an average function of the various agents across sectors (eg hospital, primary care, community care) operating on their own production functions (which may differ within and across settings) with their own levels of inputs. The overall average production function therefore may not be particularly representative of the production functions of any of the heterogeneous agents involved. Nevertheless, as (Triplett 2011) argues, the general conceptual framework is a useful starting point.

<Insert **Error! Reference source not found.** here>

For initial purposes of exposition, we assume a single output type and single input type. Following (Coelli and Prasada Rao 1998), let x_t, x_{t+1} and z_t, z_{t+1} represent the observed quantities of outputs and inputs produced in time periods t and t-1 respectively. We can represent these levels diagrammatically: in **Error! Reference source not found.** there has been an input increase from z_t to z_{t+1} and a corresponding change in observed outputs from x_t to x_{t+1}. From this information, it is possible to calculate productivity growth, given by:

$$\Delta TFP = \frac{x_{t+1}/x_t}{z_{t+1}/z_t} \qquad (1.1)$$

This measure, ΔTFP, is termed Total Factor Productivity growth. Now suppose that the production technologies in these two periods are represented by functions $f_{t+1}(z)$ and $f_t(z)$. Knowledge of each period's production function allows us to identify three components of changes in productivity over time. These are due to (1) shifts in the underlying production technology, represented by the movement of the production function; (2) changes in the scale of operation, identified by changes in input use; and (3) changes in efficiency.

If the production of output incorporated a degree of inefficiency, the output function can be represented by:

$$x_t = \lambda_t f_t(z_t)\ ,\ \text{where } 0 \leq \lambda_t \leq 1 \qquad (1.2)$$

Where λ_t is an inefficiency term for time period t, with $\lambda_t < 1$ capturing the degree of inefficiency. Substituting 1.2 into 1.1 and rearranging gives 1.3 which expresses TFP in terms of inputs, production functions and efficiency.

$$\Delta TFP = \frac{\lambda_{t+1}}{\lambda_t} * \frac{f_{t+1}(z_{t+1})/z_{t+1}}{f_t(z_t)/z_t} \quad (1.3)$$

Now suppose that there is a change in the scale of production over time. This change can be represented as proportion κ of the base periods input use, such that $z_{t+1} = \kappa z_t$. If the production function is homogenous of degree $\varepsilon(t+1)$, then 1.3 becomes

$$\Delta TFP = \frac{\lambda_{t+1}}{\lambda_t} * \frac{f_{t+1}(\kappa z_t)/\kappa z_t}{f_t(z)/z_t} \quad (1.4)$$

Which can be simplified to

$$\Delta TFP = \frac{\lambda_{t+1}}{\lambda_t} * \kappa^{\varepsilon(t+1)-1} * \frac{f_{t+1}(z_t)}{f_t(z_t)} \quad (1.5)$$

(Coelli and Prasada Rao 1998) describe 1.5 as a complete decomposition of ΔTFP where: the first ratio represents changes in technical efficiency; the second term represents the impact of a change in scale of inputs; and the final part represents technical change.

Returning to figure 1, it is clear that all three components are involved. Under the period t production function, it would not have been possible to achieve output x_{t+1}. It is feasible in period t+1 because of the shift in the production function from $f_t(z)$ to $f_{t+1}(z)$. In addition to the technological change, the observed level of output x_{t+1} is achieved because of a reduction in technical inefficiency with $\lambda_{t+1} < \lambda_t$. Finally, the increase in observed output is influenced by changes in the scale of operation, with inputs having increased. This particular depiction of the production function exhibits diminishing returns to scale. Consequently, if there had been no reduction in inefficiency, observed output in t+1 would have proportionately less than the increase in inputs.

(Coelli and Prasada Rao 1998) make several points about this formulation:

1. Without knowledge of each period's production function, even for the simplest of cases it is not possible to interpret ΔTFP as a measure of pure technical change.

2. Consequently Constant Returns to Scale are often assumed. This collapses the central term of equation 1.5, $\kappa^{\varepsilon(t+1)-1}$, to 1. There is some evidence to suggest that an assumption of CRS is tenable among NHS commissioners (Bojke, Gravelle et al. 2001) and NHS Foundation Trust hospitals and NHS hospital trusts (Castelli, Street, Verzulli and Ward, 2013 - manuscript).

3. Further, with just two year-on-year observations it is not possible to disentangle the efficiency and technical change components of ΔTFP. This leads to a further typical assumption that … '*In much of the work on the economic-theoretic approach to productivity index numbers, assumptions are made to the effect that there are no technical inefficiencies.* This assumption does not appear tenable in the face of evidence of variation in various measures of performance (Busse, Geissler et al. 2013)

Thus, because both the underlying production function and inefficiency are unobserved, it is conventional to assume that observed productivity growth is solely a function of technical change.

3.2 Index Numbers

In calculating ΔTFP for the health system, it is necessary to combine the multitude of outputs and inputs into single measures. This requires construction of output and input indices, such that total factor productivity growth is calculated by dividing an index of output growth by an index of input growth:

$$\Delta TFP = [I/Z] - 1 \qquad (1)$$

Where I is the index of output growth and Z is the index of input growth. In order to estimate total factor productivity, it is therefore necessary to correctly define and measure the output and input indices.

3.2.1 Output index

Quantification of health care output is a challenge because patients have varied health care requirements and receive very different packages of care. To address this, it is necessary to classify patients into reasonably homogenous output groupings. In order to aggregate these output groups into a single index, some means of assessing their relative value is required. Relative values are based on costs, with scaling adjustments made for changes in quality in the belief that costs are not a true reflection of social value. Our output index takes the general Laspeyres form:

$$I^{cq} = \frac{\sum_{j=1}^{J} x_{jt+1} c_{jt} \left[\frac{v_j q_{jt+1}}{v_j q_{jt}} \right]}{\sum_{j=1}^{J} x_{jt} c_{jt}}$$

We define x_j as the number of patients who have output type *j*, where *j*=1…*J*; c_{jt} indicates the cost of output *j*; q_j represents a unit of quality for output *j*, and v_j is the value of this unit of quality; and *t* indicates time. Our measures of quality include inpatient and outpatient waiting times, survival rates following hospitalisation, and blood pressure management in primary care.

3.2.2 Input index

Inputs into the health care system consist of labour, intermediate goods and capital. Growth in the use of these factors of production can be calculated directly or indirectly. A direct measure of input growth can be calculated when data on the volume and price of inputs are available such that input growth is measured as:

$$Z^D = \frac{\sum_{n=1}^{N} z_{nt+1} \omega_{nt}}{\sum_{n=1}^{N} z_{nt} \omega_{nt}}$$

Where z_{nt} is the volume of input of type *n* at time *t* and ω_{nt} is the price of input type *n* at time *t*.

This information can also be used to construct indices of price growth, holding input levels constant. A Paasche price index would measure price changes as:

$$P^D = \frac{\sum_{n=1}^{N} z_{nt}\omega_{nt+1}}{\sum_{n=1}^{N} z_{nt}\omega_{nt}}$$

Rarely are volume data available. It is common, therefore, to calculate input growth using expenditure data. Growth in total expenditure E_1 from one period to the next is driven by changes in both the volume and price of inputs. To isolate the volume effect it is necessary to convert 'nominal' monetary values into 'constant' expenditure using a deflator π which can be derived from a Paasche price index. This deflator reflects the underlying trend in prices for the inputs in question, such that $\omega_{nt} = \pi_n \omega_{nt+1}$. If these are available, the input growth index can be specified as:

$$Z^{Ind} = \frac{\pi_n E_{nt+1}}{E_{nt}} = \frac{\sum_{n=1}^{N} z_{nt+1}\pi_n\omega_{nt+1}}{\sum_{n=1}^{N} z_{nt}\omega_{nt}} = \frac{\sum_{n=1}^{N} z_{nt+1}\omega_{nt}}{\sum_{n=1}^{N} z_{nt}\omega_{nt}} = Z^D$$

4. Data

4.1 Outputs

Table 2 provides a summary of data sources used for the time period 1998/99 – 2011/12 to populate our NHS output index.

<Insert Table 1 here>

Insofar as data allow, our NHS output index captures all activities provided to NHS patients by either NHS organisations or independent sector organisations. Data for both the amount of hospital activity and for the measures of the quality of elective, non-elective and mental health care delivered in hospitals is derived from the Hospital Episode Statistics (HES). Activity performed in all NHS settings, other than hospitals and primary care, are captured in the Reference Cost returns (RC). In particular, RC data cover activity conducted in outpatient and accident and emergency departments, mental health and community care settings, and diagnostic facilities. RC also contain information on unit costs for these activities, including activity performed in hospitals.

Comprehensive data on primary care consultations are unavailable. In their absence, nationally representative survey data have been used instead. Data about the quality of primary care activity are obtained from the Quality and Outcomes Framework (QOF), which reports disease prevalence and achievement in reducing blood pressure for patients with coronary heart disease, transient ischaemic attacks or stroke and hypertension. Community prescribing data are taken from the Prescription Cost Analysis (PCA), supplied by the Prescription Pricing Authority.

There have been substantial changes over time in some of the source datasets, in particular the roll-out of Reference Costs has allowed coverage to become progressively more comprehensive. However, the disbanding of Primary Care Trusts appears to have led to a recent reduction in capture of some data. There have been regular periodic changes in the way that activities are defined and

categorised. Our output index is designed to accommodate these categorisation changes insofar as is possible (Castelli, Laudicella et al. 2011).

4.2 Inputs

We group factors of production into three broad categories, namely labour inputs; intermediate inputs, such as drugs and clinical supplies; and utilisation of capital, including buildings and equipment with an asset life of more than one year.

<Insert Table 2 here>

For each of these categories we use different, usually multiple, data sources. The summary of datasets used in different periods is presented in Table 2. From these data we construct a comprehensive index of input growth to quantify the amount of inputs used in the production of health care provided to NHS patients.

In the period from 1998/9 to 2011/2 the availability and coverage of different datasets has changed dramatically. For instance, there have been improvements in the way that data about NHS staff are collected, with it being possible to replace the 'snapshot' Workforce Census with data from the 'real-time' Electronic Staff Record (ESR) in 2007/8, the ESR providing information on number and earnings of staff categorised to more than 400 different occupational groups, rather than the 28 groups used in the Workforce Census.

Changes in datasets have not always been for the better in terms of calculation of input growth. Part of the attraction of FT status is that such Trusts are not obliged to submit yearly financial returns. Instead, we have had to rely on their less detailed annual reports for information about their expenditure.

5. Results

5.1 Outputs

Figure 2 demonstrates significant increase in NHS activity over the past 14 years. There have been steady increases in the number of people treated in hospital, with the measure of hospital allowing for changes in casemix and in the quality of their care. Hospital output has increased by 68% over the full period. There has been a 24% growth in primary care activity, though the rate of increase has been subject to some year-on-year variation. There have been year-on-year increases in prescribing, with overall growth amounting to 126%. Although there has been a 130% increase in outpatient attendances, there have been periods in which the amount of activity has fallen, notably in 2005/6-2006/7 and 2010/1-2011/2. There have also been increases in mental health activity, though a major change in classification of these activities (with the introduction of mental health "clusters") means that it has not yet been possible to report the full series. Prior to 2004/5, community care activity was not reported in a form that allowed it be included in productivity estimates prior to this date (ref RP43), but growth in such activity increased by 13% between 2004/5 and 2010/11. The apparent recent reduction may be due to failure of the organisations providing this care that to capture all the relevant data previously reported by Primary Care Trusts. This may also explain the apparent recent reduction in outpatient activity.

Whilst PbR set incentives for hospitals to increase activity, it also gave PCTs the financial means to substitute hospital care for services provided in primary or community care settings. If these services in the community provided the same health outcome as those in the hospital sector, but did so a lower price, then this would represent an efficiency gain. However, because of the cost-weighting assumptions required in constructing the productivity index, then this efficiency gain would be lost in construction. The switch from a high to a low cost *alternative* would translate to a lower valued output.

However, even casual inspection of the volumes of activity in different settings suggests little evidence of wholesale substitution. Between 2004/5 and 2010/1, there was a 13% increase in activity in community care activity. But the growth rate of elective activity has been much faster: with a 40% increase in the number of elective patients treated in hospital (6.4m to 8.9m patients). This is unlikely to be due solely to increased need of the population. If it were, elective and non-elective activity would have increased by similar proportions: non-elective activity increased by just 23% (5.7m to 7.1m patients).

<Insert Figure 2 here>

Quality improvements are captured by hospital survival rates, inpatient and outpatient waiting times, and blood pressure control in primary care. As shown in figure 3, survival rates 30 days after hospital discharge have improved year-on-year over the full period. In 1998/9, 99.29% of elective patients were alive 30-days post-discharge; this proportion had increased to 99.78% in 2011/2. The survival rate for non-electives was 94.72% in 1998/9; it had improved to 96.12% in 2011/2.

<Insert Figure 3 here>

Waiting times have varied over time. The time between GP referral and inpatient admission for elective patients remained fairly stable between 1998/9 and 2003/4, whether measured at the mean or 80th percentile of the waiting time distribution. Waiting times then fell progressively, reaching their lowest level in 2008/9. The period of stability and subsequent reductions mirror changes in waiting time policy, with the original focus on reducing the overall *number* on the waiting list being replaced by target waiting *times* which were strengthened progressively. In 2008/9, the emphasis on targets was softened, and waiting times have increased.

There are no corresponding direct targets for outpatient waiting times. However although outpatient waiting times were unavailable prior to 1999/00, they have fallen from 64 days to 24 days in 2009/10, after which the form in which they were reported was discontinued. Since 2007/8 we have calculated outpatient waits from the HES outpatients dataset. Where these data overlap with the official series, the trends are similar. Mean outpatient waiting times have been fairly stable since 2007/8.

<Insert Figure 5 here>

Measures of the quality of primary care quality were incorporated into the productivity series from 2004/5, based on data captured as part of the Quality and Outcomes Framework (Derbyshire, Zerdevas et al. 2007). If disease management in primary care is improving over time, the supposition

is that this will be reflected in reduced blood pressure for an increasing proportion of patients with CHD, stroke and hypertension. Figure 5 shows that substantial improvements in blood pressure monitoring were recorded up to 2006/7, after which progress has stabilised. This is possibly a response to the ceiling thresholds which curtails financial rewards once a target level had been reached.

5.2 Inputs

Figure 6 reports trends in expenditure on NHS labour, and trends in the volume and price (wage) of labour. As can be seen, increases in expenditure have been driven mainly by increases in wages, which have risen by 76% between 1998/9 and 2011/2. Changes in the number of NHS staff have been less pronounced, increasing by 24%.

In 2009 we undertook a fundamental review of how to measure and incorporate inputs into the productivity indices (ref RP43), including comparison of the use of expenditure and workforce data to measure the contribution of NHS labour. These alternative data sources imply similar growth in labour input, as shown by comparing the series from 2003/4. In 2007/8, we switched from using the Workforce Census to the Electronic Staff Record. From 2007/8 onwards, trends in labour input growth derived from expenditure and ESR data track each other closely.

<Insert Figure 6 here>

<Insert Figure 7 here>

There have been substantial year-on-year fluctuations in the use of agency staff. While there has been increased use of agency staff over the full series, there have been periods of retrenchment, notably between 2003/4 and 2006/7, this coinciding with the period in which the hospital sector was struggling to reduce deficits. As might be expected, in times of austerity, savings are first made in agency staff before any decrease in regularly employed NHS staff.

The 2009 review of inputs also led to a revision in the way that capital items were identified and apportioned over time in order to measure capital utilisation. In recognition of this, Figure 8 presents a broken series, each component rebased to 1998/9 and to 2003/4. The use of intermediate inputs has increased by virtually the same proportion year-on-year. In contrast, capital utilisation has been subject to greater year-on-year variation, with increased utilisation up to 2007/8, after which utilisation has declined.

The impact of PFI in the series is quite subtle. Our new capital series is intended to capture capital utilisation at each time period. If a hospital had engaged in building a new wing through traditional financing this would probably have shown up as a large expense on new capital and we would have had to apply assumptions about the consumption of that upfront expenditure over time. With PFI however, the initial capital outlay is by a separate private entity and would not show up as a capital expenditure at all in our data series. What we would observe and record is the PFI payments over time that the hospital makes to the private company, these probably being a better reflection of capitalisation at each time period than the alternative reporting format.

<Insert Figure 8 here>

5. 3 Productivity

The table below reports our estimates of year-on-year and overall output, input and productivity growth, extracted from various published reports that we have produced. Most of these reports include alternative measures, often as sensitivity analyses to assess the implications of different ways of formulating the indices or of the choice of data with which to populate them. The estimates reported in the table are our preferred measures.

<Insert Table 3 here>

While year-on-year productivity growth rates are sometimes positive and sometimes negative, looking over a succession of years reveals that output growth and input growth track each closely. Between 1998/9 and 2011/2 NHS output increased by 79% while inputs increased by 78%. The early period of the series, up to around 2003/4, was subject to negative productivity growth, with Figure Figure 10 suggesting that the NHS was operating under decreasing returns to scale. The downward trend was mainly due to big increases in inputs, with average input growth being 5.9%. In contrast, output growth increased by around 4.6% a year.

Since then, productivity growth has tended toward being positive. Annual input growth averaged 3.1% between 2005/06-2011/2, with average growth in outputs being sustained at around 4.5%. Positive productivity growth in 2004/5-2005/6 was followed by a period of constant returns to scale up to 2009/10. Productivity growth has since been positive.

6. Discussion

When discussing and evaluation a certain policy change, policy advisors tend to measure and acknowledge only the direct effects of that particular measure. There were many policy changes affecting the healthcare sector in the period from 1998/9 to 2011/2. Some of the policies may affect the same segments (ie. waiting times, labour inputs) and it is difficult to disentangle the effect of a specific change and this has not been the purpose of this paper. The purpose is, however, to look at productivity of the entire NHS and uncover any possible shifts in the productivity trends and suggest possible reasons for them.

Changes in data sources are a drawback in any analysis covering long period of time. In our wish to measure inputs and outputs as comprehensively as possible, we use several datasets which are not consistent throughout the years. As well as the data sources, also the quality of data changed considerably over time. It is hard to determine to which extent these changes contributed to our input, output and productivity indices. That said, because we construct pairwise indices our year-to-year growth indices are like-with-like comparisons.

Over the whole time series the overall picture is that growth in inputs has been broadly matched by growth in outputs. However, there has been a systematic difference in productivity growth over the early part of the period and the latter part. In the period 1998/9 to 2005/6, growth in outputs

increasingly lagged growth in inputs. This was most apparent in 2003/4 to 2005/06 when input increases were partly driven by new contract arrangements, but any incentives to improve performance had not yet occurred. From 2005/6 onwards the direct targets and incentives created by PbR, FT status, waiting time targets and other initiatives appear to have led to output growth outstripping input growth. The last two years of the series, when austerity has been the watchword, have witnessed sharp increases in productivity.

Whilst it is difficult to precisely pin down the changes in productivity to specific policies, the observed patterns do conform to the more predictable elements of the policies and we observe different patterns across settings: for instance, the increase in hospital activity which might be expected as a result of PbR and waiting time targets which is not matched by similar increases in primary care activity. Similarly primary care quality has responded up to the point at which it has been rewarded (and no further) and waiting times have decreased until they are no longer considered a priority.

Acknowledgements

The project was funded by the Department of Health in England as part of a programme of policy research at the Centre for Health Economics, University of York. The views expressed are those of the authors and may not reflect those of the funder.

Agenda for Change Project Team (2004). Agenda for Change – Final Agreement. Leeds, Department of Health.

Bloor, K., N. Freemantle, et al. (2012). "Trends in consultant clinical activity and the effect of the 2003 contract change: retrospective analysis of secondary data." Journal of the Royal Society of Medicine **105**: 472 - 479.

Bojke, C., H. Gravelle, et al. (2001). "Is bigger better for primary care groups and trusts?" BMJ **322**: 599-602.

Busse, R., A. Geissler, et al. (2013). "Diagnosis Related Groups in Europe: Moving Towards Transparency, Efficiency, and Quality in Hospitals?" British Medical Journal **346**.

Castelli, A., M. Laudicella, et al. (2011). "Getting Out What We Put In: Productivity of the English National Health Service." Health Economics, Policy and Law **6**(3): 313-335.

Coelli, T. and D. S. Prasada Rao (1998). An Introduction to Efficiency and Productivity Analysis, Springer.

Department of Health (2004). The NHS improvement plan: Putting people at the heart of public services. London, Department of Health.

Derbyshire, K., P. Zerdevas, et al. (2007). Further developments in measuring quality adjusted healthcare output. Leeds, Department of Health.

Mannion, R., G. Marini, et al. (2008). "Implementing payment by results in the English NHS: changing incentives and the role of information." J Health Organ Manag **22**(1): 79-88.

Sahoul, J., A. Stafford, et al. (2011). "NHS CAPITAL INVESTMENT AND PFI: FROM CENTRAL RESPONSIBILITY TO LOCAL AFFORDABILITY." Financial Accountability & Management **27**(1).

Talbot-Smith, A. and A. Pollock (2006). The New NHS: A Guide: A Guide to Its Funding, Organisation and Accountability, Routledge.

Triplett, J. E. (2011). Health System Productivity. The Oxford Handbook of Health Economics. P. C. Smith and S. Glied, Oxford : Oxford University Press.

Labour Government is elected

Coalition Government is elected

Comprehensive Spending Review (CSR) 2000
27 per cent increase in nominal NHS funding over three years

Comprehensive Spending Review (CSR) 2004
21 per cent increase in funding over three years

Comprehensive Spending Review (CSR) 2007
22 per cent increase in funding over three years

Comprehensive Spending Review (CSR) 2010
11 per cent increase in funding over four years

Publication the White Paper: Our healthier nation.

Publication of the NHS plan: a plan for investment, a plan for reform

Health Secretary Patricia Hewitt announces £620m deficit for 2005/6

Publication the White Paper: Building the National Care Service.

Publication the White Paper: Liberating the NHS.

The National Health Service (Private Finance Act 1997) becomes law introducing the **Private Finance Initiative (PFI)** to fund capital investment in the NHS

The creation of **Foundation Trust (FT)** hospitals is announced

Introduction new **GP Contract and QOF**

Introduction new **Hospital Consultant Contract**

Introduction of **Agenda for Change**

19971 19981 19991 20002 20012 20022 20032 20042 20052 20062 20072 20082 20092 20102 20112 20122

Payment by Results (PbR) introduced for 15 HRGs**Payment by Results (PbR)** introduced for 15

Target of max **18 weeks wait** from referral to treatment is introducedTarget of max **18**

PbR extended to all hospitals and includes A&E, outpatients and emergency admissions**PbR** extended to all hospitals and

Best Practice Tariffs introduced in PbR**Best Practice Tariffs** introduced in PbR

NHS structure

NHS Direct is established

New **Primary Care Groups (PCGs)** are established

District Health Authorities are replaced by **Strategic Health Authorities (SHAs)** and **Primary Care Trusts (PCTs)**

The first **FTs** hospital are established

SHAs reduced from 28 to 10
PCTs reduced from 303 to 152

The Health and Social Care Bill proposes to (a) increase GPs' influence on commissioning, (b) increase competition and (c) abolish SHAs and PCTs

Table 1: Summary of NHS output data sources, 1998/99 – 2011/12

Output type	Activity source	Cost source	Year	Quality	Notes
Elective and day cases	Hospital Episode Statistics	Reference Costs	1998/99 – 2011/12	· 30 day survival; · Health outcomes; · Waiting times.	
Non-elective	Hospital Episode Statistics	Reference Costs	1998/99 – 2011/12	· 30 day survival; · Health outcomes.	
Outpatient	Reference Costs	Reference Costs	1998/99 – 2011/12	· average waiting time (weeks).[1]	*Classification of outpatient activity changed over time, so has the waiting time measure.*
Hospital and non-admitted Mental Health	Hospital Episode Statistics and Reference Costs	Reference Costs	1998/99 – 2011/12	· 30 day survival; · Health outcomes; · Waiting times.	*Quality adjustment applied to Hospital Mental Health activity only.*
Community Care	Reference Cost)	Reference Costs	1998/99 – 2011/12	· n/a	
Other NHS activity	Reference Costs	Reference Costs	1998/99 – 2011/12	· n/a	
Primary Care	General Household Survey	PSSRU's Unit Costs of Health and Social Care	1998/99 – 2003/04	· n/a	
	QResearch GP Patient Survey		2003/04 – 2008/09 2008/09 – 2011/12	· quality adjustors (prevalence and achievement rates in reducing blood pressure for patients with CHD, stroke and hypertension) (Quality Outcome Framework)	
Prescribing	Prescription Pricing Authority (PPA)	Prescription Pricing Authority (PPA)	1998/99 – 2011/12	· n/a	*Switched from counting total volume of prescriptions to items of prescription in 2008/09.*

[1] Outpatient waiting times up until and including 2009/10 were provided by the Department of Health, National Inpatient and Outpatient waiting times, England. From 2009/10 onwards, we use the HES Outpatient Minimum Dataset.

Table 2: Summary of Input data sources, 1998/99 – 2011/12

Input type	Activity source	Year	Deflator	Notes
Direct Labour	Workforce Census	1998/99-2006/07	HSHC Pay Index	For years 1998/99-2003/04 data was quality adjusted
	Electronic Staff Record (iView)	2007/08-2011/12	HSHC Pay Index	
Indirect Labour	Trusts financial returns	1998/99-2011/12	HSHC Pay Index	
	Foundation Trust financial Returns	2003/04-2011/12	HSHC Pay Index	
	SHA, PCTs financial returns	2003/04-2011/12	HSHC Pay Index	
Intermediates	Trusts financial returns	1998/99-2011/12	NHS prices index	
	Foundation Trust financial Returns	2003/04-2011/12	NHS prices index	
	SHA, PCTs financial returns	2003/04-2011/12	NHS prices index	
Capital	Trusts financial returns	1998/99-2011/12	ONS MM17 capital deflators	
	SHA, PCTs financial returns	2003/04-2011/12		
	Foundation Trust financial Returns	2003/04-2011/12		
Prescription Drugs	Number of prescription	1998/99-2006/07	FHS deflator	
	Number of items	2007/08-2011/12	CHE pharmacy price index	

Table 3: Input, Output and Productivity Growth

	QA Output	Output growth Index	Input	Input growth Index	QA Productivity	Productivity growth index
		1		1		1
1998/99-1999/00	2.22%	1.02	5.07%	1.05	-2.71%	0.97
1999/00-2000/01	2.26%	1.05	1.55%	1.07	0.69%	0.98
2000/01-2001/02	3.74%	1.08	6.10%	1.13	-2.22%	0.96
2001/02-2002/03	5.78%	1.15	7.06%	1.21	-1.19%	0.95
2002/03-2003/04	4.93%	1.20	7.63%	1.30	-2.51%	0.92
2003/04-2004/05	6.44%	1.28	6.50%	1.39	-0.05%	0.92
2004/05-2005/06	7.11%	1.37	7.19%	1.49	-0.07%	0.92
2005/06-2006/07	6.50%	1.46	1.92%	1.52	4.50%	0.96
2006/07-2007/08	3.66%	1.51	3.88%	1.58	-0.21%	0.96
2007/08-2008/09	5.73%	1.60	4.23%	1.64	1.44%	0.97
2008/09-2009/10	4.11%	1.67	5.43%	1.73	-1.25%	0.96
2009/10-2010/11	4.57%	1.74	1.33%	1.76	3.21%	0.99
2010/11-2011/12	2.51%	1.79	1.53%	1.78	0.96%	1.00

Notes: Figures up to 2002/03-2003/04 (light blue shaded) are based on Research Paper 6: output growth table 9.3, page 182; productivity growth table 9.6, page 186. Figures for 2003/04-2004/05 are taken from Research Paper 47: input growth table 5.3, page 28, column 5; output growth table 5.4, page 29, final row; productivity growth table 5.6, page 30, column 2; Figures for 2004/5-2010/11 (light green shaded) are taken from Research Paper 87: output growth table 11, page 15; input growth table 14, page 23, column 1; productivity growth table 15, column 1. The final row of figures is yet to be finalised.

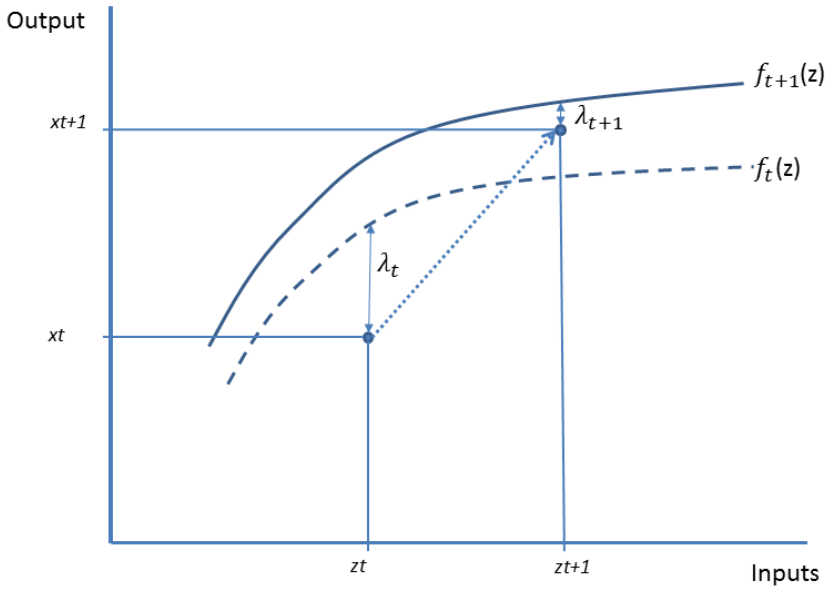

Figure 1: Theoretical presentation of production function

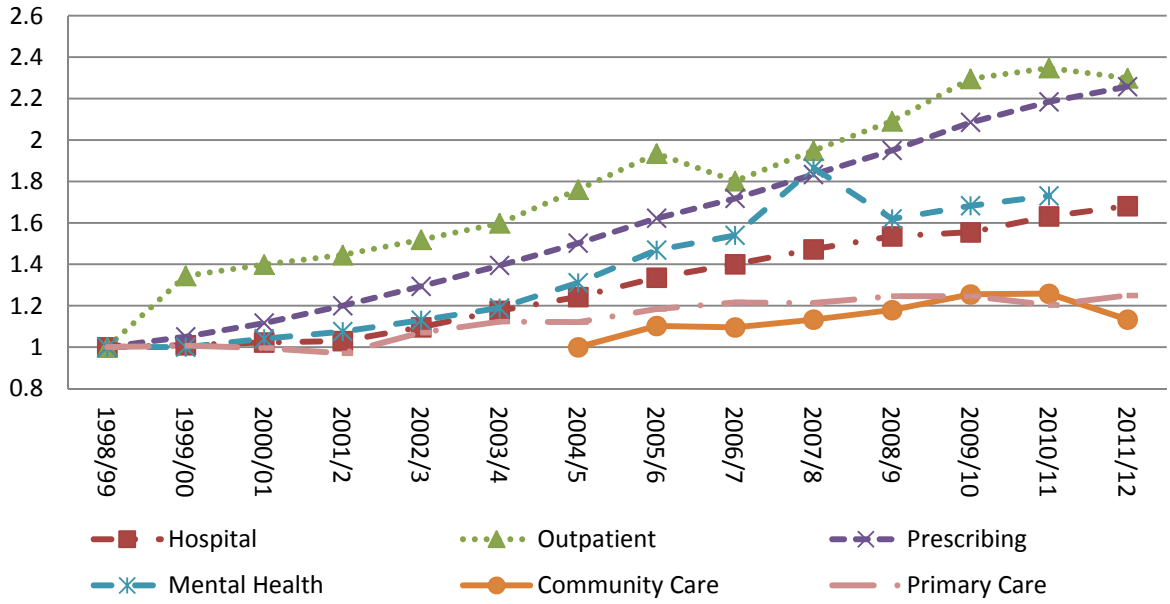

Figure 2: Output growth by sectors; rebased to 1998/99.

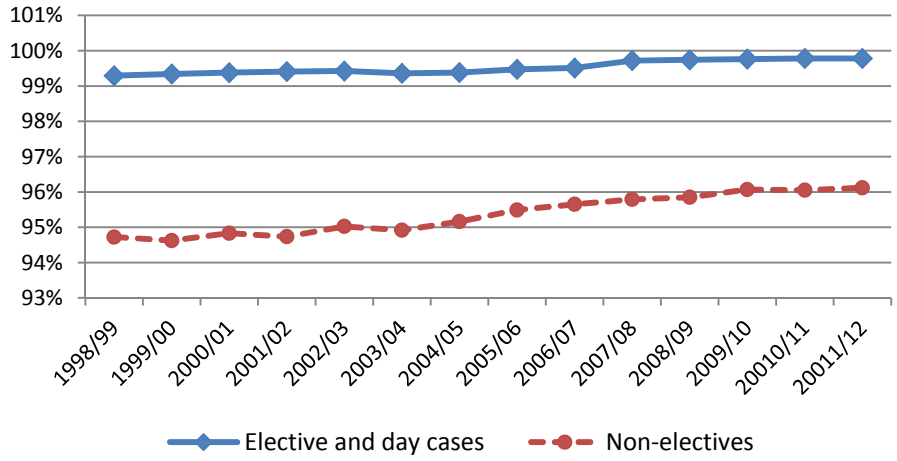

Figure 3: 30-day survival rates, by type of hospital activity

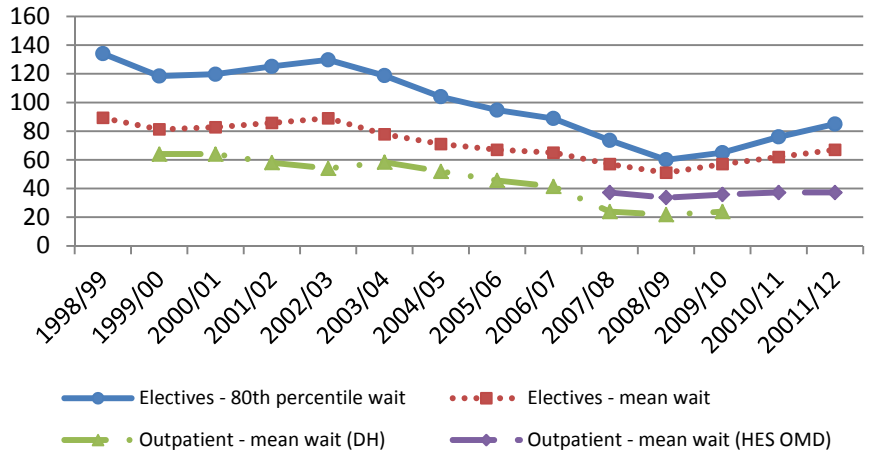

Figure 4: 80th percentile and mean waiting times, 1998/99 – 2011/12

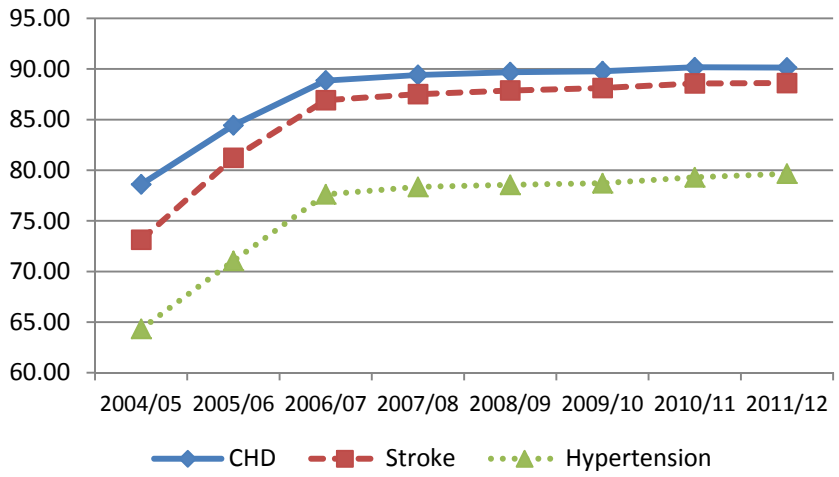

Figure 5: QOF achievement for proportion with low blood pressure

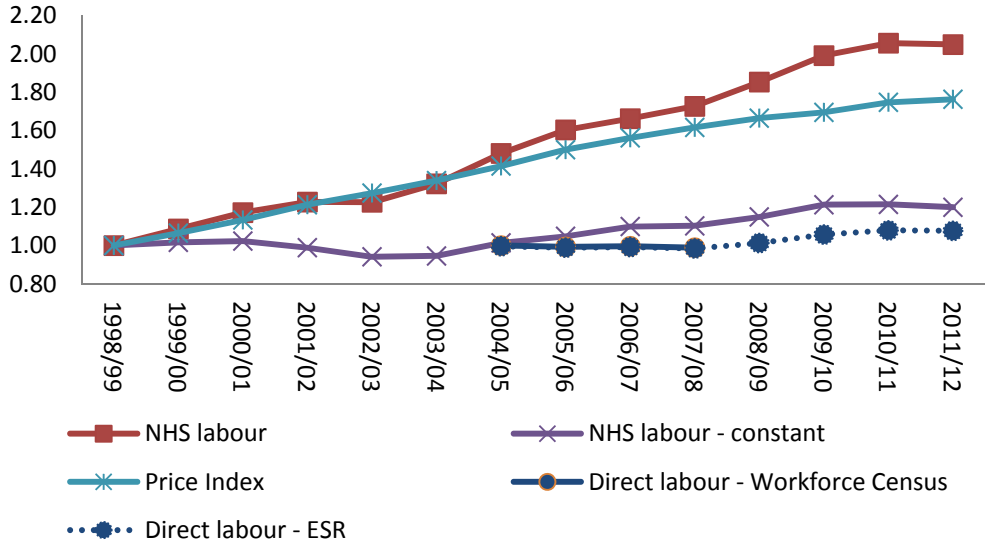

Figure 6: Labour growth

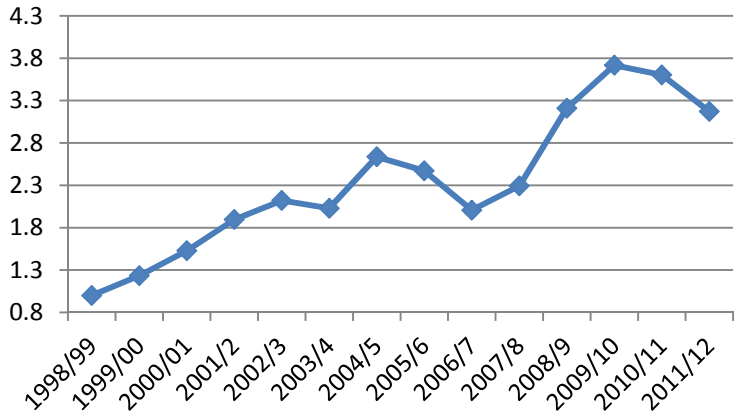

Figure 7: Agency growth

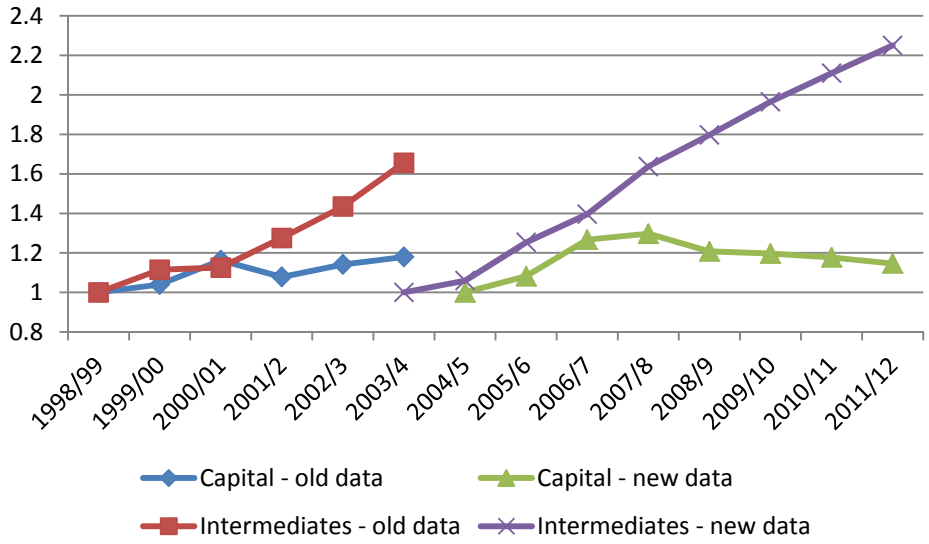

Figure 8: Growth in capital and intermediates

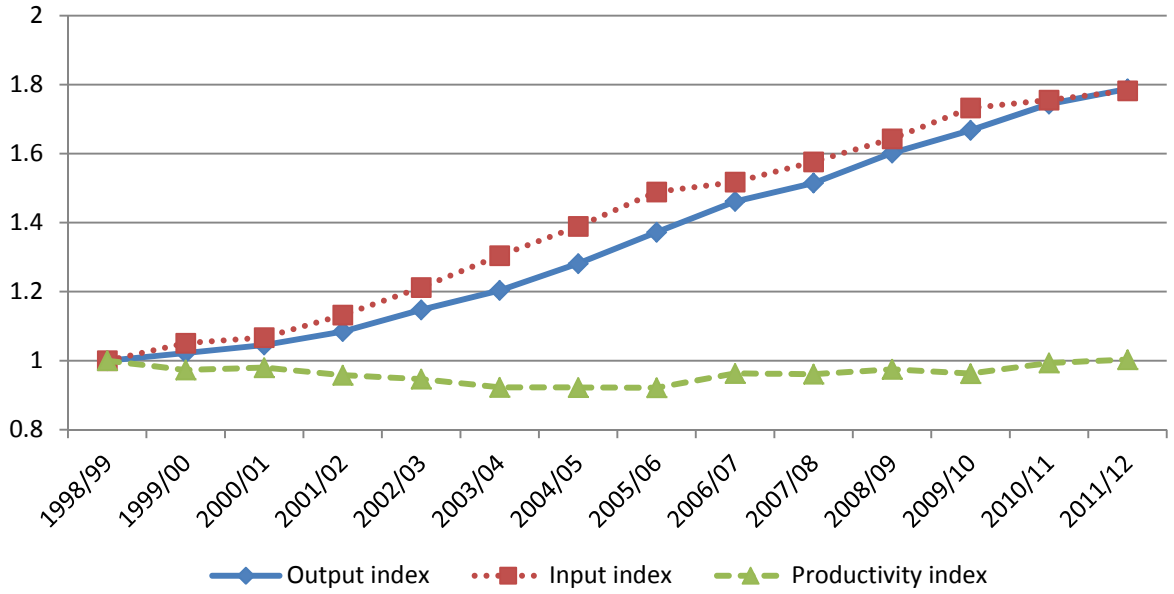

Figure 9: Graph of input, output and productivity index

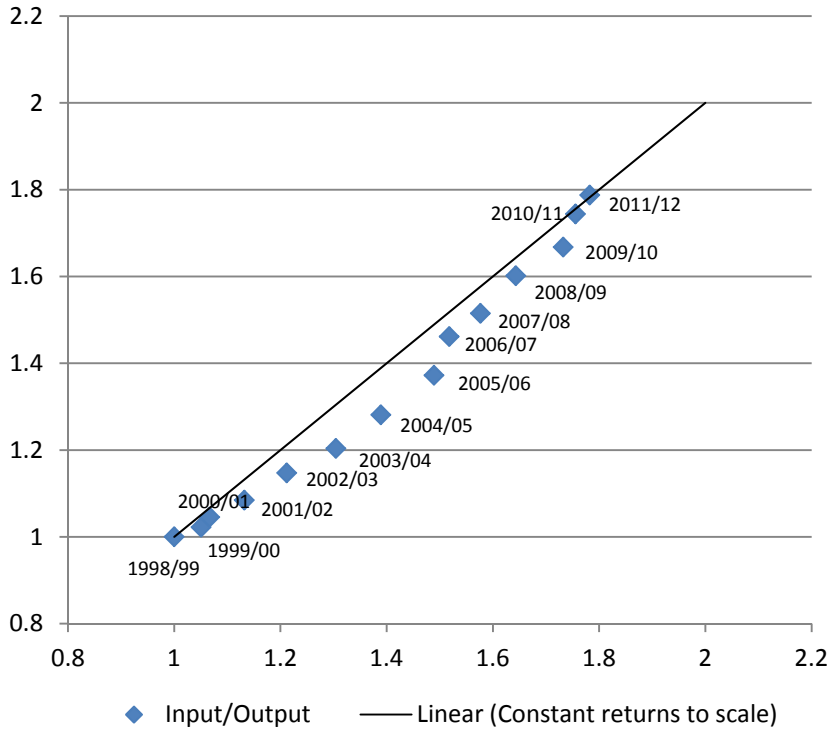

Figure 10: Simple plot of the input and output index