International Comparisons of Performance in the Provision of Public Services:

Outcome based measures for Education.

Mary O'Mahony and Philip Andrew Stevens

National Institute of Economic and Social Research, London

January 2004

Abstract

This paper considers the measurement of performance in public service provision in an international context. It first sets out the measurement issues in general terms. The paper then applies these methods to estimate labour productivity in the education sector comparing the UK experience with that in the US from the mid 1990s. The results suggest reasonable labour productivity growth in UK education over this time period and show the UK outperforming the US.

The authors acknowledge funding for this research came from the 'Evidence Based Policy Fund' with contributions from ONS/DfES/DoH/HM-Treasury.

1 Introduction

This paper considers the issue of measuring performance in the provision of public services. Frequently studies use volume of output measures such as number of students educated or numbers of medical interventions. An earlier paper (O'Mahony, Stevens and Stokes, 2002) set out the arguments for and against using information on final outcomes to measure the services provided, such as increases in average years of life due to medical interventions or lifetime earnings arising from education. It concluded that there were strong theoretical arguments in favour of using outcome information rather than relying solely on outputs. This was based on the argument that the lack of market valuations (prices), exacerbated by incomplete information, suggests final outcomes may yield a more accurate measure of the effectiveness of the services provided. But the paper also pointed out that in practice it may be very difficult to implement an outcome based measure, in particular to adequately take account of factors that affect outcomes but are extraneous to the service provider. Examples are lifestyle changes that increase life expectancy or technological changes that increase the effectiveness of particular types of skilled labour.

The purpose of this paper is a first attempt to implement an outcome approach and to compare performance across countries in the provision of public services. The application chosen is education, arguably the easiest service to measure in an international context. Nevertheless this paper will show that there are a number of practical problems that need to be resolved. The paper's primary concern is to derive a measure of relative productivity performance for the entire education sector, to complement research carried comparing performance in services in the private sector (O'Mahony and deBoer, 2001). International comparisons are an important benchmark in examining performance and may often yield more insights than comparisons across time for a single country alone. The former is useful in evaluating the extent to which different systems of provision impact on performance whereas the latter is most useful in examining the impact of within country changes. There is also an issue relating to expectations on the magnitudes of increases in service provision or productivity over time. O'Mahony and DeBoer (2001) show that growth rates vary considerably across sectors, with manufacturing showing on average more than 2% growth in output per hour worked between 1989 and 1999, whereas services sectors such as financial and business services achieved no more than 1% in the same period. Hence International comparisons can also aid in benchmarking expectations on what is achievable.

The next section briefly sets out the measurement issues in general terms and presents a measure of performance to be applied to the education sector. The main body of the paper then applies this method to the education sector with a view to estimating labour productivity growth rates comparing the UK and the US in the second half of the 1990s. It begins by describing the output volume data available for the UK, underlying the estimates of output growth in the UK National Accounts. This section also examines information from an outcome measure, test score results, that is frequently employed in evaluating performance across producing units. We argue that this measure is sensitive to weights on the various test score results and so does not as yet provide a practical alternative to output measures. Methods employed to measure earnings outcomes are set out in section four and regression results for both countries are presented. Section five employs the resulting estimates on lifetime earnings to estimate outcome based measures of aggregate education services for the UK. It first considers the use of these outcomes as weights for various types of education in deriving an aggregate measure. It then examines methods that might be used to incorporate changes in the effectiveness of education across time. Finally this section presents measures of UK labour input and labour productivity that suggest growth rates close to those achieved in the economy as a whole. Section six presents the outcome, input and labour productivity results for the US and then compares results in the two countries. This shows the UK outperforming the US in all years. In contrast, relative labour productivity growth in the total economy shows the US outperforming the UK in most years of the time period considered here. Finally section seven concludes with an outline of extensions of the education application in future research.

2 Outputs, outcomes and productivity: Measurement issues.

This section sets out definitions and methods to measure output and inputs in public services. To start assume at time t a particular sector J provides n services, Y^J_i, using k inputs, X^J_i. Examples of the former are education at various levels (primary, secondary, university etc.) or types of medical interventions and of the latter are teaching staff (primary and secondary teachers, teaching assistants and university lecturers) and school buildings or medical staff (doctors, nurses etc.) and medical equipment. Let Q_i denote the quantity of services produced, e.g. number of pupil hours by type or number of operations. A measure of the growth in service provision in an aggregate sector can be calculated as:

$$D(t)Y_t^J = \sum_i w(t)_i^J \, D(t)Q_{i,t}^J \tag{1}$$

Where the operator D(t) denotes the log rate of change, D(t)Y = ln(Yt) – ln(Yt-1) and w are weights on the i types of services provided. Similarly productivity growth in sector J can be calculated as:

$$D(t)P_t^J = D(t)Y_t^J - \sum_k s(t)_k^J \, D(t)X_{k,t}^J \tag{2}$$

Where s are the weights on the k inputs. In what follows we first concentrate on measuring service provision (1) and then go on to discuss additional issues relating to the measurement of input growth.

In measuring output growth in private market services the weights w in equation (1) are estimated by the share of each individual service in the total value of output produced, i.e. (for simplicity omitting time subscripts) by:

$$w_i^J = \frac{p_i^j Q_i^J}{\sum_i p_i^J Q_i^J} \tag{3}$$

If (3) is averaged across time periods t and t-1, and substituted into (1) then we have the commonly employed Tornqvist index of output growth. In addition changes in the quality of services across time can be incorporated by replacing Q in (3) by volume measures in effectiveness units. In practice this is achieved by estimating Q using deflated values, with quality adjusted prices replacing actual market prices in the deflation.

If a service is provided by the public sector, however, market prices do not exist and therefore we lack a measure of the marginal benefit to consumers of the service provided. In the past measures have been employed whereby the cost of producing service i is used as a weight. Suppose each service uses only one input unique to that service, with unit costs c_i, then these alternative weights are given by:

$$x_i^J = \frac{c_i^j X_i^J}{\sum_i c_i^J X_i^J} \tag{4}$$

Extending (4) to multiple inputs for each output is straightforward, with the total cost terms replaced by a sum across inputs used. Weights such as (4) have been used, for example, in the cost weighted activity index (CWAI) measure of the output of the UK health sector calculated by the Department of Health. An alternative approach to cost weights is to place value judgements on the relative merits of different types of services. However these 'judgmental' weights are likely to be controversial at best and open to abuse at worst.

The main problem with using cost shares in publicly provided services is that there is no market mechanism that ensures that the marginal cost of providing the service equals the marginal benefit to consumers and this may result in significant divergence between the two. For example a medical intervention may be very expensive but yield little by way of increases in life expectancy or quality of life. Cost weighting gives such treatments an unjustifiably high weight. In addition it is difficult to incorporate quality adjustments in the cost share approach, i.e. there is no natural equivalent to deflating by quality adjusted prices. There is no doubt that quality aspects are important in public service provision, most notably in medical care since improvements in medical procedures are substantial and so it is important to include quality adjustments. Technological innovation may lower the cost of providing a particular service while at the same time rendering it more effective. The use of cost share weights would then lead to a lowering of the impact of this service on aggregate growth but in reality it should have a greater weight. Innovations that allow for out-patient treatments for particular ailments at considerably lower costs have been quite common in health provision and often these are more effective than the hospital treatments they replaced (e.g. the example of treatment of depression or cataracts- see discussions in papers in Cutler and Berndt, 2001). These observations underly the recent disquiet with the CWAI measure for health.

However these problems extend beyond the lack of markets since much of the literature suggests that, even when services are privately provided, the market price may not reflect the 'true' benefit to consumers if there are information asymmetries between providers and consumers. In health care the consumer has an inadequate basis for making a judgement on whether medical interventions are worthwhile, i.e. for making informed choices among both providers and types of interventions. In addition there is an argument that health care insurance places a wedge between the producer and consumer with consequent moral hazard problems. Triplett (2001) argues that the most important difference between services such as health care and general market services such as car repair is that in the latter case the consumer can sell or scrap the car but this is not possible for the consumer of health services (human repair). We have social norms that prevent consumers making this decision. Thus even in private health services, information provided by the market is inadequate to allow consumers to judge the quality of the service they are consuming.

The lack of market prices and arguments on information asymmetries suggest that it may be more useful to measure performance using outcomes rather than outputs. Letting O_i denote the outcome from the provision of service i, the simplest measure is merely to sum across outputs, assuming they are measured in consistent units as discussed further below. The growth in this 'total outcome' measure is given by:

$$D(t)O^J = D(t)\sum_i O_i^J \tag{5}$$

The first problem with implementing (5) is that outcomes are a function of many factors other than the direct provision of a service. For example in health care we can write Health outcomes (HO) as follows:

HO = H (medical interventions, diet, lifestyle, environment, genetic factors, etc.)

Since medical interventions are one of a number of contributing inputs to the production of health, it is natural to measure the contribution of medical intervention by its incremental contribution. Similarly education outcomes depend on a range of background variables including the social and ethnic mix of the population as well as the inherent ability of the students being educated.

In general terms we can write service i outcomes in the form:

$$O_i^J = \Phi(Q_i^J, Z_i^J) \tag{6}$$

where Z are extraneous or background influences. The incremental contribution of the service provided to outcomes is then given by the partial derivative with respect to Q in equation (6):

Incremental contribution = δ (O)/ δ (Q)

other influences constant.

There are a number of methods that could be used to measure the incremental contributions. At a detailed level we could focus on particular types of services, controlling for population differences. Thus the disease based approach in the OECD project on Age Related Diseases (ARD) is an example whereby researchers consider detailed medical records for a subset of the population (the elderly) – see also papers in Cutler and Berndt, 2001. Ultimately such a detailed approach is likely to provide the most robust findings but is very intensive in research time. Alternatively we can use a regression based approach by regressing outcomes such as earnings or life expectancy on the service provided and a range of control variables. Regression methods are employed in the education example in section 3 of this paper. Thus in principle we can estimate a measure O*, adjusting for the influence of extraneous factors, and substitute this into equation (5).

In practice, however, it may not be possible to adjust for all background influences in a single step, in particular if the factors that affect outcome levels are very different from those that affect growth rates and information on both come from different data sources. In this case an alternative way of proceeding is to use a two step 'outcome flow' method whereby information on outcomes are first used to calculate the weights in equation (1) and then the result is adjusted for outcome growth. Thus these weights are given by:

$$wo_i^J = \frac{O*_i^J}{\sum_i O_i^{*J}} \tag{7}$$

where O* are outcome values having adjusted for the influence of background variables. Substituting (7) into (1) gives the outcome flow measure as:

$$D(t)YO_t^J = \sum_i wo(t)_i^J \, D(t)Q_{i,t}^J \tag{8}$$

Although useful as a devise to weight the quantity of services, equation (8) does not allow outcomes to change over time at a differential rate to changes in quantities, i.e. it does not take account of changes in effectiveness through time. Therefore it may be necessary to adjust equation (8) by adding a term involving some additional growth in outcomes. Letting Q* denote outputs measured in effectiveness units, then ideally we wish to estimate the following:

$$D(t)YO'^J_t = \sum_i wo(t)_i^J \, D(t)Q*_{i,t}^J \tag{9}$$

The application to education discusses a number of methods of incorporating adjustments for increases in effectiveness, one based on an age cohort analysis and a second based on an adjustment for the impact of education on long term economic growth. Nevertheless this step remains the most difficult to incorporate in practice.

In order to implement (8), (9) (and in practice equation (5)) all outcomes need to be translated into some common metric. Otherwise we would have to include additional weights in defining O* in equation (7) and hence would be essentially back to where we started with equation (1). One approach would be to translate all outcomes into monetary values, adjusting for general inflation, and this we see as probably the best way forward. Thus in education the outcome would be lifetime earnings arising from participation in education and in health this would be the values of additional years of life through medical interventions. Note that the absolute monetary value placed on the outcomes does not feature in equation (7) since the weights are outcome shares. Rather what matters is the relative impact on outcomes of the services provided.

The next section sets out an application of the 'outcome flow' approach to education. Before doing so however we need to consider the input side of the productivity equation (2). Here there is much less difficulty since providers of public services must bid for inputs in the same market as private firms. Hence the wages paid to inputs can be used to derive cost share weights so that aggregate input is derived as a Tornqvist index of individual inputs. In the remainder of this paper we only consider labour input and labour productivity measures. Future extensions will also incorporate capital inputs.

3 Education outputs and outcomes

The remainder of this paper considers the practical application of the outcome approach to international comparisons for the education sector. This is confined to a comparison between the UK and the US. This service was chosen since, at the outset, it appeared that the measurement issues were more transparent than in more complex areas such as health or social services. The primary quantity of output measure, numbers of pupils educated, is relatively easy to measure with plentiful data available to compare across countries. An obvious candidate as an outcome measure is test score results but its use leads to some difficulties, as discussed further below. There is also a clearly defined outcome measure, taking the lead from Jorgenson and Fraumeni (1991), i.e. the impact on lifetime earnings arising from education. In measuring the latter it is possible to draw on a vast academic literature to set out the estimation issues and survey data can be employed to estimate returns to education at each level. Nevertheless a large number of measurement problems arise even in this relatively simple application. The purpose of this section is to set out clearly the issues involved.

3.1 Output and Outcome Measurement.

In this analysis we will consider three measures of the output of the education sector, a volume measure, and two 'quality adjusted' outcome measures based on test scores and earnings, respectively. The starting point for each measure is a Tornqvist chain linked index, of the form of equation (1), based on the growth in pupils/students in each educational level between time periods t and t-1, given by:

$$\Delta Q_t = \sum_i \omega_{i,t} \Delta \ln(PUP_{i,t}) \tag{10}$$

where PUP is the number of pupils in education, i is the level of education, ω is a weighting factor and Δ is the first difference operator. By setting a base year equal to 100, the growth rates in (10) can be used to construct an index of the output of the education sector. This general framework allows us to measure the annual flow of services of the education sector.

In this section we first consider the three measures using UK data. Following this we present estimates for the US and then compare the results for the two countries.

3.1.1 Volume of output

The simplest volume measure is to set ω equal to the shares of type i pupils in total pupils. An alternative frequently employed volume measure is to weight each type in total pupil/student numbers by the share of total expenditure on level i education. But as argued in section 2, this approach is best avoided. Pupil shares ensures that the output measures are independent of input changes which is an important consideration when we consider productivity growth.

In this preliminary analysis we confine attention to the period 1994 to 2001 when data exist for all measures. The quantity of output measure employed at ONS is changes in pupil hours. They have in fact assumed that the hours each pupil is taught per annum is fixed so 'pupil hours' is actually measured as the total number of pupils being taught (full-time equivalent). This gives an indication of the change in the volume of education output. Following ONS we divide the UK education system into seven levels, nursery, primary, secondary, further education (below degree level), undergraduate, postgraduate and special schools.

Table 1 shows index numbers from 1994 to 2001 for these categories for the UK. The final two rows show shares of pupil numbers in the base year 1995 and annual average growth rates across the period. In total pupil/student numbers increased by 2.4 per cent per annum with the largest increases in the two higher education categories. Higher education represented about 11% of pupil/student numbers in 1995 – by 2001 this had risen to 13%.

Large increases were also found in nursery, which however represented only a small proportion of the total. Over this period pupils aged 16 plus grew faster than those up to age 15 in secondary schools. Further education also shows above average increases.

Table 1 Pupil/Student Numbers, UK, 1994-2001

	Nursery	Primary	Secondary		Further	Higher Education		Special	Total
			up to age 15	age 16+	Education	Undergrad	Postgrad		
1994	99.0	98.7	97.8	99.2	93.6	93.8	92.0	98.9	96.8
1995	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
1996	99.4	101.6	98.7	105.8	97.6	102.9	99.8	98.9	100.3
1997	98.9	102.4	99.7	112.7	93.3	106.4	99.3	99.0	100.5
1998	99.2	102.9	100.1	114.3	92.0	118.0	119.4	99.2	102.0
1999	98.2	102.7	102.1	114.2	91.3	115.8	111.8	98.5	101.9
2000	121.1	102.0	102.3	114.5	109.0	117.5	115.8	97.6	105.8
2001	142.8	101.3	103.9	115.5	114.5	123.4	129.1	96.6	107.9
SH95	0.48	38.76	23.76	3.88	21.33	9.03	1.88	0.88	100.00
G	5.22	0.37	0.85	2.17	2.88	3.92	4.84	-0.34	1.55

SH95 = share in total pupils 1995, G = annual average growth rate, 1994-2001

Thus the greatest percent increases have been achieved at the ‘higher quality’ end of the distribution. A straightforward pupil weighted index does not capture this quality differential.

3.1.2 Quality adjusted output: test score outcomes.

An obvious candidate to construct a quality adjusted output measure is to incorporate the results from test scores into the analysis. Table 2 shows a range of test score measures at different education levels which in theory could be applied to the volume measures in Table 1. All measures show large increases over the period with those in primary and GCSEs dominating.

Table 2 Test Scores: UK education, selected levels

	KS2:[1] Level 4 or greater	GCSE: percent 5 or more A*-C	A-levels:[2] Percent 3 or more	HE: percent 1st and 2.1
Per cent of pupils/students				
1994	62.5	46.6	66.1	47.2
1995	62.5	47.8	68.9	47.5
1996	62.5	48.4	70.0	47.8
1997	62.5	49.3	69.0	48.2
1998	62.0	50.9	69.4	49.7
1999	70.0	53.1	70.0	50.2
2000	73.5	54.7	86.2	51.1
2001	73.0	55.7	74.2	52.9
Index 1995=1.00				
1994	1.00	0.98	0.96	0.99
1995	1.00	1.00	1.00	1.00
1996	1.00	1.01	1.02	1.01
1997	1.00	1.03	1.00	1.01
1998	0.99	1.07	1.01	1.05
1999	1.12	1.11	1.02	1.06
2000	1.18	1.14	1.06	1.08
2001	1.17	1.17	1.08	1.11

Notes:

1. *Average English and Maths. Note primary test scores not available before 1997 so these were assumed constant up to then;*
2. *As % attempting A-levels.*

In principle it should be possible to utilise the information in Table 2 with the volume measures in Table 1 to arrive at a quality adjusted measure. in order to do so we need to impute a weight to pupils/students who achieve the threshold level relative to those who do not reach this level (with the latter normalised to equal one). Thus for each education level i, we compute a pupil effectiveness index

$$\mathrm{PUPE}_i = \alpha_i\, \mathrm{PUP}_i\, \lambda_i + (1- \alpha_i)\, \mathrm{PUP}_i \tag{11}$$

Where α_i is the percent of level *i* pupils achieving the threshold score and λ_i is the effectiveness ratio. Summing across the *i* levels gives a 'quality adjusted' alternative to equation (10):

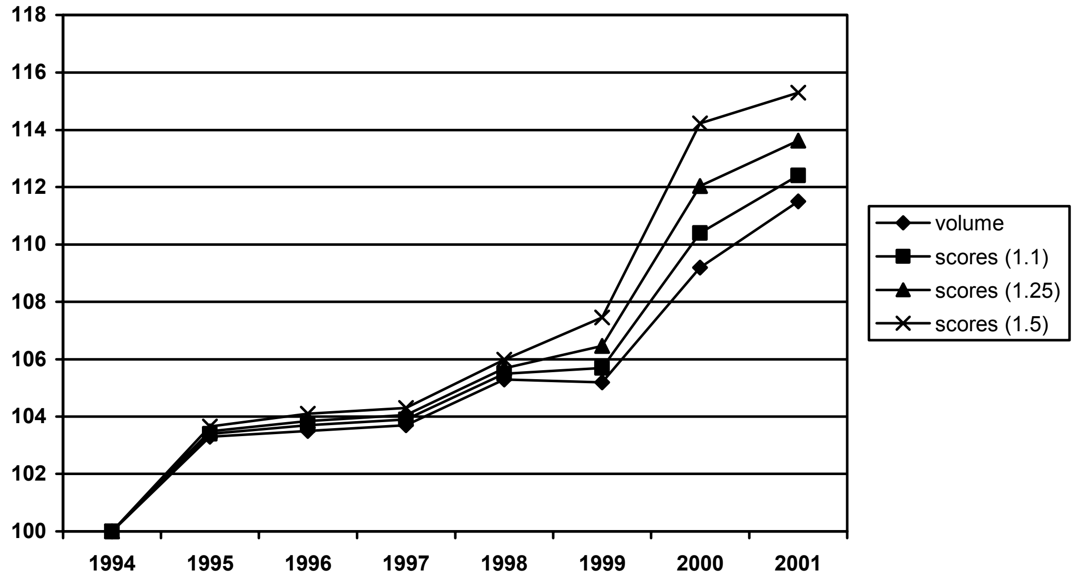
$$\Delta QE_t = \sum_i \omega^e{}_{i,t} \Delta \ln(PUPE_{i,t}) \tag{12}$$

where ω^e is the share of type *i* effective pupils in PUPE (averaged across period *t* and *t*-1 as in (1)).

The problem in using test scores is that there is no basis on which to impute the effectiveness ratios, λ_i . In addition it is necessary to impute a value to education levels not covered by test score statistics. The most reasonable assumption is to use the closest equivalent category for omitted ones (primary scores for nursery and special schools and A-levels for further education). Assuming λ is the same across education levels, calculations based on (3) are shown in Figure 1 for three variants together with the volume measure for comparison purposes. These assume pupils/students achieving the threshold values are 10%, 25% and 50%, respectively, 'more effective' than pupils who do not reach this level. This shows that the results are sensitive to the weights employed but all three show faster growth than a crude volume measure, varying from higher growth of about 0.1% p.a. to 0.5% p.a. depending on the weights used.

In reality we would expect the effectiveness measures, λ, to vary by type of education received but by how much is difficult to gauge. One possibility would be to use information on relative earnings. However detailed estimates of the impact on earnings of an additional GCSE or A-level or comparisons of graduate earnings by grade of degree awarded are not readily available.

Figure 1 Quality Adjusted Indexes based on Test Scores for UK Education.



There are a number of additional problems in attempting a calculation of this kind. First the results are sensitive to the cut-off point in each indicator. For example using percent of pupils aged 15 with 1 or more GCSE and the lowest 10% 'effectiveness' weight would lower the overall growth by about 0.4% per annum since this indicator grows much less rapidly than the indicator in Table 2. Further problems arise when there is a suspicion of 'grade inflation' so that increases in the scores may not reflect any true improvement. This is a concern with both secondary and higher education levels. Again this could be dealt with if there were detailed data on earnings – if improved scores are primarily due to grade inflation then the market will not remunerate workers with improved scores. Against this it may be the case that some tests are set up so that there is a general tendency for diminishing returns to set in at some stage. Thus in primary education the tests are set up so that pupils are required to pass some (time invariant) threshold. Increasing effort may be required to get pupils at the lowest end of the ability to pass this threshold. Finally achieving test score results may be subject to extraneous influences outside the education sector, such as effort put in by parents. This is less important when considering changes over time than when comparing across pupils or schools at a point in time, but nonetheless remains a concern.

4 Earnings outcomes.

An alternative to the use of test scores is to use information from earnings in the marketplace to weight achievements at each education level. This section first considers the estimation of wage premiums to education, controlling for other influences on earnings. We then use this information to derive an earnings outcome based approach to measure education provision.

4.1 Estimating education wage premiums and lifetime earnings

In this study we will concentrate on the private financial return to education as a measure of education outcomes. Whilst we do not downplay the importance of the social return to education investment, such research is not possible within the confines of this particular project. The effect of education on labour market outcomes has a number of dimensions. Most important of these is the wage an individual can expect to earn with a given level of education. Another factor, which has an indirect impact on earnings, is the probability that an individual will be able to find a job in the first place. There are essentially three potential labour market states an individual can find themselves in their lifetime: (*i*) in work; (*ii*) unemployed; and (*iii*) economically-inactive. Difficulties arise in valuing the non-pecuniary aspects of each of these states (non-wage benefits of working; the value of spare time when unemployed or inactive). Jorgenson and Fraumeni (1991) take one extreme view of this when they attempt to calculate the value of each hour spent in work and leisure. They argue that individuals are free to choose their hours and will do so such that the marginal value of work and leisure are equal. The implication of this for working individuals is that each hour of leisure (except that spent sleeping) is worth the same in dollar terms as those in work. This has a number of difficulties for working people. It assumes that workers are indeed free to choose the hours they work, or at least able to make a trade off between working hours and wages. A difficult assumption to sustain is that the higher paid have a better quality of non-working life than the less well paid.

The first step is to estimate the impact of education on earnings controlling for extraneous influences. In our analysis of the outcome on earnings we will employ a standard Mincerian human capital earning function[1]. In the standard model estimates the log of earnings as a function of years of schooling and a second or more-order polynomial of experience. For example:

$$\ln(Y) = \alpha_0 + \alpha_1 s + \alpha_{21} e + \alpha_{22} e^2 + \sum_k \beta_k C_k + \varepsilon_i \tag{13}$$

where

Y = income,
s = years of schooling,
e = experience (years in employment),
C = a vector of control variables
ε = is an error term $\varepsilon \sim N(0,\sigma)$.

There are a number of issues relating to the estimation of such equations. The first is the question of whether years of schooling represent the correct measure of schooling. This may be valid in countries like the US, Card (1999) argues, but less so in countries like Germany and France, which have multiple education streams. For our purpose, it is important to link expenditure, via outputs to outcomes. The public sector in general allocates funds not to an extra year of education, but rather to particular types of education:

i.e. to specific levels (e.g. primary), for particular qualifications (e.g. the new AS level), or a particular initiative (e.g. targeting mathematical skills). Therefore, for this study it is more appropriate to replace the *s* term in (13) with terms for the particular level of education experienced and/or qualification obtained. That is

$$\ln(Y)=\alpha_0+\sum_l \alpha_{1l}q_l+\alpha_{21}e+\alpha_{22}e^2+\sum_k \beta_k C_k \tag{14}$$

where *q* represent *l* levels of education.

In the UK the breakdown is as follows:

	Qualification level	*Variable name*
1.	No qualifications	*NOQUAL*
2.	Secondary education up to GCSE	*GCSE*
3.	Secondary education up to A-Level	*ALEVEL*
4.	Trade Apprenticeships	*TRADAPP*
5.	Further Education qualification	*FE*
6.	Higher education – Undergraduate	*HE_UG*
7.	Higher education – Postgraduate	*HE_PG*

The baseline category in the regressions is no qualifications. For completeness trade apprenticeships are included as a separate category although these do not feature in our education outputs.

In the US it is:

[1] The study of the private returns to education has a long history. David Card provides a useful survey of the

	Qualification level	*Variable name*
1.	Less than 11^{th} Grade	-
2.	11^{th} Grade	*GRADE11*
3.	12^{th} Grade, but no Diploma	*GRADE12N*
4.	12^{th} Grade, High School Diploma, GED	*GRADE12D*
5.	Some college but no degree	*SOMECOL*
6.	Associate degree	*ASSDEG*
7.	Undergraduate degree	*UG*
8.	Postgraduate of professional degree	*PG*

The baseline category in the regressions is education to less than 11th Grade.

One important issue to bear in mind when considering the effect of education on earnings is the fact that we only observe wages for those individuals who are in work. We may not observe the wages of others for two main reasons. The first is that individuals cannot find wage at a level that is high enough to entice them into work. Because of this, we will not observe the lower end of the wage distribution and so estimates of the effect of education on income will be biased upwards. If wages are increasing in education, this bias will be worse at lower levels of education, since fewer individuals with lower levels of education will be offered a wage. This will, at least in part, be offset if better educated individuals have higher reservation wages. Tied up with this is the fact that some individuals may leave the labour market for other reasons, such as childbearing. The second reason why we do not observe an individual's wage is because there is no work available at any wage. That is, the individual is unemployed.

Because of the potential for our estimates to be biased, we employ a 'Heckman correction/selection' methodology in estimating wages (Heckman, 1976). In this model equation (14) is modified to account for the fact that the dependent variable in the earnings

causal effect of education on earnings in his Handbook of Labor Economics chapter (Card, 1999).

regression is only observed if a secondary inequality is satisfied (the 'selection equation'). That is, the dependent variable in equation (14) for individual i is only observed if

$$\gamma z_i + \upsilon_i > 0 \tag{15}$$

where the error term $\upsilon \sim N(0,1)$ and corr(ε, υ) = ρ.

When $\rho \neq 0$, a standard regression of equation (14) will yield biased results. The Heckman selection model provides consistent, asymptotically efficient estimates for all the parameters in such models. In the results we report the Wald test of independence of the selection and earnings equations, i.e. the likelihood ratio test that ρ = 0. In addition to reporting ρ, we also report the selectivity effect, $\lambda = \rho\sigma$, as well as its standard error.

Before we continue, we must note that there are two additional potential biases in OLS estimates of the returns to education. The first is due to an omitted variable measuring the innate ability of an individual, the second is that family background may also affect an individual's educational attainment. There is a long history of using instrumental variables (see Card, 1999, for a survey). These instrumental variables analyses tend to find higher returns to education. However, Dearden suggests that conventional OLS estimates of the returns to education can generally be relied upon for policy decisions after estimating models which take account of individual ability and parental influence on education. Therefore, in order to keep this work as transparent as possible we do not follow an instrumental variables approach.

4.2 The Influence of Education on Economic Activity

Another way in which education has an impact on the labour market experience of individuals is via its affect on economic activity. Not only are those with higher levels of education likely to attract higher wages, they are also less likely to be unemployed and may also be less likely to be economically inactive (Stevens, 2003). Therefore, when attempting to measure the influence of education on lifetime earnings, it is also important to assess the impact on economic activity. Therefore we estimate a multinomial-logistic model of the probability of an individual being in one of three labour market states (employed, unemployed, inactive).

The probability that person *i* finds themselves in any one of these mutually exclusive states is given by

$$\Pr(Z_i = j) = \frac{e^{X_i \beta_j}}{\sum_{j=1}^{3} e^{X_i \beta_j}} \tag{16}$$

where X_i is a vector of explanatory variables, β is a vector of coefficients to be estimated, j = 0, …, 3 are the potential outcomes (0 = employment, 1 = unemployment, 2 = economically inactive). In order to remove the indeterminacy of the model, we normalise by setting $\beta_0 = 0$. That is, the probabilities we calculate are the probability of the particular outcome relative to being employed. The probabilities of each outcome are, therefore,

$$\begin{aligned} \Pr(Z_i = j) &= \frac{e^{X_i \beta_j}}{1 + \sum_{j=1}^{2} e^{X_i \beta_j}} \quad \text{for } j = 1,2 \\ \Pr(Z_i = 0) &= \frac{1}{1 + \sum_{j=1}^{2} e^{X_i \beta_j}} \quad \text{for } j = 0. \end{aligned} \tag{17}$$

4.3 The Total Effect of Education on Lifetime Earnings

The total effect of education on lifetime earnings is the product of the wages an individual might expect to earn if working and the probability of not working. Not earning a wage influences our estimation of lifetime earnings in two ways. First, as we have seen, it may bias our estimates of the determinants of earnings if there are any systematic differences between individuals for whom we have earnings data and those for whom we do not. Second, people without work will earn nothing, or at least have a much lower level of income, such as unemployment benefit or insurance. In what follows, we assume that unemployed people in the UK earn the basic rate of unemployment benefit, and those in the US earn unemployment insurance equal to half of their earnings (implicitly this involves two simplifying assumptions: that they are not unemployed for periods longer than 26 weeks in a 'benefit year' and that they do not usually earn more than the threshold). Therefore, the total effect of education on earnings is

$$E_q = W_q P_q(emp) + UP_q(unemp) + 0 \times \left[1 - P_q(unemp) - P_q(emp)\right] \tag{18}$$

where E_q = the total earnings associated with education level q, W_q is the predicted effect on wages from the earnings regression, $P_q(emp)$ is the probability of being employed from the activity regression, U is the income the individual would obtain if they were unemployed, P_q is the probability of being unemployed from the activity regression. Thus equation (6) says that expected earnings are the sum of the chances of being employed multiplied by the wages that would be earned, the unemployment benefit multiplied by the chances of being unemployed plus the chances of being inactive multiplied by zero (i.e. we assume that individuals gain nothing financially from economic activity).

4.3.1 Results for the UK

The results for the UK are presented in Table 4 and Table 5. These are based on data from Summer 1996 to Spring 1997. Table 4 presents the results of the earnings estimation. The Wald test of independence is significantly different from zero (χ^2 = 1860.77), clearly justifying the Heckman selection model.

Earnings are increasing in experience and educational qualifications for both men and women, although the effect of experience is decreasing because of the non-linearity in the specification. We can see the importance of examining the effects of qualifications on wages rather than simply years of education by the fact than the returns to A-Levels and FE are very different. Although, the coefficients on both in our earnings equation are statistically different from that of on the GCSE variable, that for FE is much lower than that for A-Levels. Those with trade apprenticeships as their highest qualification typically earn less than those with FE qualifications. However, there has been a considerable decline in the numbers undertaking trade apprenticeships and so reflects these structural changes in the labour market.

One explanation for the difference between the returns to further education and those to A-Levels is an unobserved ability bias. It is likely that those students who enter further education are have lower levels of innate ability, and certainly lower GCSE scores, than those who take A-Levels. In order for these estimates of returns to truly represent additional earnings power engendered by further and sixth-form education we would need more detailed information on GCSEs or some measure of innate ability[2].

Turning to the activity equations in Table 5 we see a similar pattern emerge. The probability of unemployment is declining in qualification level. Again further education has a smaller effect in reducing the likelihood of unemployment than A-Levels.

The results are similar for inactivity, with the probability of an individual being economically inactive declining with education. Unlike the results for unemployment, those who undertake further education are less likely than those with A-Levels to be economically inactive.

---

[2] Although Cawley, Heckman and Vytlacil (1998) argue that measures of cognitive ability and schooling are so highly correlated as to make separating their effects impossible.

We can compare our results to other work by converting our coefficients into a 'per year' equivalent, to give an estimate of the rate of return. We do this by subtracting from the coefficient for a particular level that for the previous level and dividing this by the years of additional schooling required for the extra qualification. For example, if we wish to consider the rate of return for a year of undergraduate study for men, we first subtract the coefficient for the return to an A-Level education from that for undergraduate studies to obtain the additional earnings due to undergraduate studies (0.49970 - 0.30196 = 0.19774). We then divide this number by the number of years it takes to complete undergraduate education (typically three) to get a rate of return for undergraduate studies of 0.065913, or 6.6%. This compares to the estimated average return to a year of schooling of 6.5% in the OLS results of Chevalier and Walker (2001) for the UK in 1995 (using the Family Expenditure Survey). Chevalier and Walker (2001) also undertake a similar estimation of the returns to qualifications using the British Household Panel Survey (although their breakdown of qualifications is different). The results for postgraduate and undergraduate degrees and A-Levels are of a similar order to ours, although their returns to GCSEs are much higher. This may be due to differences in specification, since their variables have different qualifications subsumed in them; they also include a number of vocational qualifications separately and do not include ethnic effects, but do include regional effects.

Table 3 Implied Rates of Return, UK

	Coefficient	*Years of schooling*	*Rate of Return*
Men			
Secondary education up to GCSE	0.03033	Same as for no qualifications	-
Secondary education up to A-Level	0.30196	GCSE+2	0.135815
Further Education qualification	0.19363	GCSE +2	0.08165
Higher education – Undergraduate	0.49970	A-level +3	0.065913
Higher education – Postgraduate	0.58566	HE UG +3	0.028653
Women			
Secondary education up to GCSE	0.00122	Same as for no qualifications	-
Secondary education up to A-Level	0.28150	GCSE+2	0.14014
Further Education qualification	0.11219	GCSE +2	0.055485
Higher education – Undergraduate	0.53961	A-level +3	0.086037
Higher education – Postgraduate	0.65474	HE UG +3	0.038377

Table 4 Earnings Equations, UK

Using Heckman Selection Method

	Men		Women	
	Earnings equation	*Selection equation*	*Earnings equation*	*Selection equation*
potexp	0.12380*** (0.00297)		0.08376*** (0.00347)	
potexp^2	-0.00469*** (0.00017)		-0.00435*** (0.00020)	
potexp^3	0.00006*** (0.00000)		0.00007*** (0.00000)	
Health problem	0.06135*** (0.01238)	-0.35686*** (0.01644)	0.21715*** (0.01543)	-0.40223*** (0.01546)
Black	-0.01065 (0.03971)	-0.31829*** (0.05313)	0.33829*** (0.04689)	-0.39813*** (0.04726)
Indian	0.12089*** (0.03839)	-0.40033*** (0.05111)	0.34497*** (0.05016)	-0.42583*** (0.04958)
Pakistani/ Bangladeshi	-0.10126** (0.04839)	-0.49762*** (0.06044)	0.70475*** (0.08179)	-1.05527*** (0.07104)
Other Asian	0.17452*** (0.06432)	-0.62626*** (0.08152)	0.49163*** (0.07419)	-0.45197*** (0.07333)
Mixed	-0.06557 (0.07070)	-0.06923 (0.09805)	0.20330** (0.08235)	-0.09121 (0.08535)
Other	0.11090 (0.11057)	-0.59353*** (0.13969)	0.49335*** (0.13752)	-0.64536*** (0.13103)
HE_PG	0.58566*** (0.02576)	0.64915*** (0.03817)	0.65474*** (0.03613)	0.61533*** (0.04014)
HE_UG	0.49970*** (0.01857)	0.47880*** (0.02517)	0.53961*** (0.02215)	0.51814*** (0.02256)
FE	0.19363*** (0.01732)	0.47966*** (0.02296)	0.11219*** (0.02033)	0.55726*** (0.02035)
TRADEAPP	0.15119*** (0.01811)	0.24832*** (0.02413)	0.09098*** (0.03373)	0.20078*** (0.03418)
ALEVEL	0.30196*** (0.02128)	0.35128*** (0.02874)	0.28150*** (0.02542)	0.35377*** (0.02569)
GCSE	0.03033* (0.01639)	0.38578*** (0.02139)	0.00122 (0.01771)	0.40445*** (0.01750)
DKQUAL	0.88053*** (0.05221)	-1.44071*** (0.05017)	1.43092*** (0.07035)	-1.33373*** (0.05633)
age		0.18443*** (0.01546)		0.02323 (0.01423)
age^2		-0.00446*** (0.00044)		0.00057 (0.00040)
age^3		0.00003*** (0.00000)		-0.00001*** (0.00000)
married		0.22894*** (0.01246)		-0.09507*** (0.00963)
Constant	8.67921*** (0.02130)	-2.45660*** (0.17126)	8.77172*** (0.02444)	-0.93397*** (0.15839)
ρ	-0.87		-0.95	
σ	0.73		0.99	
λ	-0.64	(0.01)	-0.94	(0.01)
χ^2	1860.77		2625.72	
$p(\chi^2)$	0.00		0.00	
Observations	40829		42856	
censored	18928		20336	

- *Standard errors in parentheses*
- ** significant at 10%; ** significant at 5%; *** significant at 1%*
- χ^2 *= Likelihood ratio test of* $\rho = 0$

Table 5 Activity equations, UK

Multinomial logit (omitted category = in employment)

	Men		Women	
	Unemp	*Inactivity*	*Unemp*	*Inactivity*
age	-0.11392** (0.05210)	-0.74551*** (0.04062)	-0.13372** (0.06331)	0.23104*** (0.02956)
age²	0.00127 (0.00152)	0.01631*** (0.00116)	0.00272 (0.00188)	-0.00903*** (0.00084)
age³	0.00000 (0.00001)	-0.00011*** (0.00001)	-0.00002 (0.00002)	0.00010*** (0.00001)
Married	-1.05186*** (0.05588)	-0.81887*** (0.04477)	-0.77075*** (0.06228)	0.02776 (0.02807)
Health problem	0.67049*** (0.05517)	2.14487*** (0.03840)	0.49158*** (0.06607)	1.15115*** (0.02799)
Black	1.11179*** (0.12473)	0.71900*** (0.13287)	1.11639*** (0.12950)	0.54280*** (0.08945)
Indian	0.52697*** (0.16022)	0.55681*** (0.12958)	0.84025*** (0.16263)	0.53165*** (0.08887)
Pakistani/ Bangladeshi	1.17498*** (0.14347)	1.09317*** (0.12757)	1.53778*** (0.20964)	2.05590*** (0.11729)
Other Asian	0.90149*** (0.23537)	1.92392*** (0.15730)	0.80831*** (0.26670)	0.90871*** (0.13029)
Mixed	0.63706** (0.25837)	0.40058* (0.23626)	0.66451** (0.26591)	0.42649*** (0.16063)
Other	1.37044*** (0.34391)	1.28737*** (0.31919)	0.96922** (0.43963)	1.16348*** (0.21604)
HE_PG	-1.72277*** (0.18140)	-1.64641*** (0.14678)	-1.41683*** (0.21590)	-2.19488*** (0.11279)
HE_UG	-1.29574*** (0.08898)	-1.18331*** (0.07266)	-1.05639*** (0.10366)	-1.70076*** (0.04953)
FE	-0.96322*** (0.06915)	-1.25020*** (0.05988)	-0.73128*** (0.08309)	-1.40473*** (0.04008)
TRADEAPP	-0.97049*** (0.08165)	-0.94141*** (0.06165)	-1.17464*** (0.20365)	-0.74257*** (0.06361)
ALEVEL	-1.29766*** (0.10228)	-0.19482*** (0.06483)	-0.89905*** (0.11117)	-0.87753*** (0.04816)
GCSE	-0.90798*** (0.06190)	-0.88895*** (0.04921)	-0.57008*** (0.06994)	-0.93709*** (0.03163)
DKQUAL	-0.69264*** (0.23880)	-0.31490* (0.18481)	-0.81897** (0.37018)	-0.59708*** (0.14245)
Constant	0.70785 (0.55139)	8.38751*** (0.43102)	0.19053 (0.65683)	-1.66531*** (0.32376)
Observations	39176	39176	41781	41781

- *Standard errors in parentheses*
- ** significant at 10%; ** significant at 5%; *** significant at 1%*

In order to test the sensitivity of our results to the choice of year, we also performed the same analysis on data from the summer 1998 to spring 1999 quarters. The results of these analyses are presented in the Appendix as Table 12 and Table 13. We can see that the results are fairly similar, with returns to university education being slightly higher and those to FE slightly lower. The figures for the returns to A-levels and GCSEs are approximately equal.

4.4 Results for the US

The results for the US are presented in Table 7 and Table 8. Again, the Wald test of independence is significantly different from zero (χ^2 = 1860.77), justifying our use of the Heckman selection model. Earnings in the US are also increasing in education and experience as we would expect *a priori*. The rate of return to the 11th and 12th grades (without achieving a diploma) are similar, at around 8.5%. Achieving a diploma has a strong positive effect on earnings, although it is unlikely that the comparison with those who achieve only 10th or 11th grade is appropriate here, since it is likely that most if not all of those who could achieve a High School Diploma continue until 12th grade and so those who drop out before 12th grade are come from a similar population to those who stay to 12th grade and do not obtain a Diploma. Likewise, the return to those attending college and obtaining an associate degree is just under 8%, whereas for those who do not obtain a degree it is actually negative, i.e. they have similar earnings to those who are only educated to 11th grade. The returns to an undergraduate degree are much higher than those to associate degrees and there is little return to postgraduate degrees over and above undergraduate study. To put these figures in perspective, Trostel, Walker and Wooley (2002) estimate the returns to a year of schooling in the US to be 12.99% and 14.66%, for men and women respectively, which is consistent with an average of our rates of return.

Table 6 Implied Rates of Return, US

	Coefficient	*Years of schooling*	*Rate of Return*
Men			
11th Grade	0.0852	<11th grade + 1	0.0852
12th Grade, but no Diploma	0.1693	11th grade + 1	0.0841
12th Grade, High School Diploma, GED	0.38718	11th grade + 1	0.30198
Some college but no degree	0.06029	12th grade + 1	-0.3269
Associate degree	0.5405	12th grade + 2	0.07666
Undergraduate degree	0.94224	Ass deg + 3	0.13391
Postgraduate of professional degree	0.9097	UG + 3	-0.0108
Women			
11th Grade	0.0852	<11th grade + 1	-0.8245
12th Grade, but no Diploma	0.1693	11th grade + 1	0.0841
12th Grade, High School Diploma, GED	0.38718	11th grade + 1	0.30198
Some college but no degree	0.06029	12th grade + 1	-0.3269
Associate degree	0.5405	12th grade + 2	0.07666
Undergraduate degree	0.94244	Ass deg + 3	0.13398
Postgraduate of professional degree	0.9097	UG + 3	-0.0109

Table 7 Earnings Equations, US

Using Heckman Selection Method

	Men		Women	
	Earnings equation	*Selection equation*	*Earnings equation*	*Selection equation*
potexp	0.07655*** (0.00148)		0.06196*** (0.00157)	
potexp^2	-0.00242*** (0.00009)		-0.00229*** (0.00010)	
potexp^3	0.00003*** (0.00000)		0.00003*** (0.00000)	
Black	-0.12484*** (0.00888)	-0.14183*** (0.01374)	0.02503*** (0.00915)	-0.08737*** (0.01185)
American Indian	-0.06522*** (0.02273)	-0.19949*** (0.03413)	0.03212 (0.02607)	-0.15235*** (0.03208)
Asian	-0.10297*** (0.01259)	-0.07081*** (0.02009)	0.08876*** (0.01478)	-0.14616*** (0.01846)
PG	0.90970*** (0.01191)	0.32221*** (0.01886)	0.77257*** (0.01532)	0.84310*** (0.01949)
UG	0.94224*** (0.01010)	0.38101*** (0.01623)	0.73544*** (0.01286)	0.75519*** (0.01559)
ASSDEG	0.54050*** (0.01218)	0.37647*** (0.02006)	0.33964*** (0.01451)	0.78249*** (0.01804)
SOMECOL	0.06029*** (0.00714)	-0.00653 (0.01154)	0.02343*** (0.00810)	0.11502*** (0.01046)
GRADE12D	0.38718*** (0.00921)	0.25422*** (0.01406)	0.18933*** (0.01191)	0.53696*** (0.01375)
GRADE12N	0.16930*** (0.02206)	0.15881*** (0.03351)	0.03173 (0.02903)	0.30448*** (0.03397)
GRADE11	0.08520*** (0.01441)	0.08564*** (0.02061)	-0.10843*** (0.01811)	0.26627*** (0.02068)
age		0.36288*** (0.00991)		0.11707*** (0.00900)
age^2		-0.00911*** (0.00028)		-0.00226*** (0.00025)
age^3		0.00007*** (0.00000)		0.00001*** (0.00000)
married		0.21723*** (0.00798)		-0.16986*** (0.00658)
Constant	5.40611*** (0.01027)	-4.08790*** (0.10712)	5.49119*** (0.01266)	-1.79879*** (0.09840)
ρ	-0.86		-0.92	
σ	0.73		0.87	
λ	-0.63	0.00	-0.79	0.00
χ^2	4828.49		7166.78	
p(χ^2)	0.00		0.00	
Observations	106448		112926	
censored	31248		40507	

- *Standard errors in parentheses*
- ** significant at 10%; ** significant at 5%; *** significant at 1%*
- χ^2 *= Likelihood ratio test of* $\rho = 0$

Table 8 Activity equations, US

Multinomial logit (omitted category = in employment)

	Men		Women	
	Unemp	*Inactivity*	*Unemp*	*Inactivity*
age	-0.17391*** (0.03915)	-0.65619*** (0.02457)	0.05100 (0.04437)	-0.05779*** (0.01854)
age2	0.00295*** (0.00113)	0.01453*** (0.00070)	-0.00228* (0.00130)	-0.00078 (0.00052)
age3	-0.00002* (0.00001)	-0.00009*** (0.00001)	0.00002* (0.00001)	0.00002*** (0.00000)
married	-0.09376*** (0.03586)	-1.05966*** (0.02503)	-0.28653*** (0.03951)	0.49200*** (0.01684)
Black	0.68408*** (0.04267)	0.73569*** (0.02888)	0.75330*** (0.04241)	0.21310*** (0.02263)
American Indian	0.77667*** (0.09888)	0.71824*** (0.07134)	0.72782*** (0.11126)	0.35778*** (0.05859)
Asian	0.05922 (0.07786)	0.57020*** (0.04644)	0.16948* (0.08884)	0.33952*** (0.03421)
PG	-0.57442*** (0.07487)	-1.67612*** (0.05828)	-1.60002*** (0.10372)	-2.08839*** (0.04113)
UG	-0.87678*** (0.06326)	-1.42183*** (0.04261)	-1.48555*** (0.07206)	-1.66849*** (0.02900)
ASDACA	-0.68924*** (0.07646)	-1.26613*** (0.05457)	-1.38702*** (0.08547)	-1.71946*** (0.03491)
ASDVOC	-0.08716** (0.04189)	0.14388*** (0.02823)	-0.36807*** (0.04816)	-0.18687*** (0.02000)
GRADE12D	-0.37080*** (0.04823)	-0.86302*** (0.03026)	-0.78252*** (0.05402)	-1.16227*** (0.02412)
GRADE12N	-0.14655 (0.10639)	-0.41424*** (0.06849)	-0.37980*** (0.12399)	-0.54299*** (0.05856)
GRADE11	-0.12270* (0.06544)	-0.27381*** (0.03927)	-0.33293*** (0.07247)	-0.49700*** (0.03531)
Constant	0.65421 (0.40902)	7.68793*** (0.25692)	-1.92146*** (0.45956)	1.55901*** (0.20028)
Observations	106448	106448	112926	112926

- *Standard errors in parentheses*
- ** significant at 10%; ** significant at 5%; *** significant at 1%*

5 Lifetime earnings and productivity: results for the UK

5.1 Lifetime earnings as weights: results for the UK

The results of the analysis in the previous section gives the lifetime earnings achievable for given levels of education and given assumptions on activity rates. The results for the UK are summarised in the Table 9 below.

Table 9 Discounted Lifetime earnings, UK.

	Discounted lifetime earnings*	
	Total	Relative to no qualifications
HE_post graduate	£228,736	1.501
HE_under graduate	£240,710	1.579
FE	£209,248	1.373
A-level	£227,941	1.495
GCSE	£185,587	1.218
No Qualifications	£152,424	1.000

** average of male and female earnings*

We then want to combine this information with the volume measure (number of pupils/students) to arrive at an annual outcome flow measure. This is achieved by translating the results in Table 9 into an incremental flow of lifetime earnings from each year spent in the education system. To do so requires some assumptions regarding the lifetime earnings within each education level. The simplest approach is to assume equal step increases for each year of schooling within the education level. For example consider the case of primary school education which typically lasts 7 years. Pupils who complete one year of primary school are deemed to earn one seventh of the primary school addition to lifetime earnings, two years earn two sevenths etc. Thus the additional output for pupils in their second year is 1/7 *LTE where LTE is (discounted) lifetime earnings from primary education. Thus in any one year, if we assume an equal distribution of pupils across the seven years, the primary school output is (1/7) LTE times the number of primary school pupils. In general for each of i education levels lasting for k years, earnings outcomes in each year are given by

$$EO_{i,t} = \sum_{i} \frac{1}{k} LTE_i PUP_{i,t} \tag{19}$$

The share of type i total earnings outcomes in total EO can then be used as weights in a Tornqvist index:

$$\Delta QO_t = \sum_i \omega^o{}_{i,t} \Delta \ln(PUP_{i,t}) \tag{20}$$

where $\omega^o_{i,t} = 0.5\left[\frac{EO_{i,t}}{\sum_i EO_{i,t}} + \frac{EO_{i,t-1}}{\sum_i EO_{i,t-1}}\right]$

The formula in equation (20) is similar to that employed by Jorgenson and Fraumeni (1991) in estimating the output of the US education sector. The main difference is that in this paper our estimates of lifetime earnings are based on regression analysis whereas Jorgenson and Fraumeni use US census data divided by sex, age and educational attainment. In addition Jorgenson and Fraumeni allow their lifetime earnings shares to vary across time whereas we base our estimates on regressions for a single year. Future extensions of this paper will attempt to incorporate time varying shares.

This formula was applied to the UK education data from 1994 to 2001.[3] Note in implementing this approach we need to impute a value for earnings if an individual had no education whatsoever. This we base on the minimum wage in both countries. The results are shown in Figure 2 together with the crude volume measure and the test score measure with effectiveness ratio set to 1.1. The lifetime earnings relative to no qualifications in Table 9 suggest weights in this range rather than 1.25 or 1.5. The three measure show similar trends in the beginning of the period but diverge towards the end with the earnings

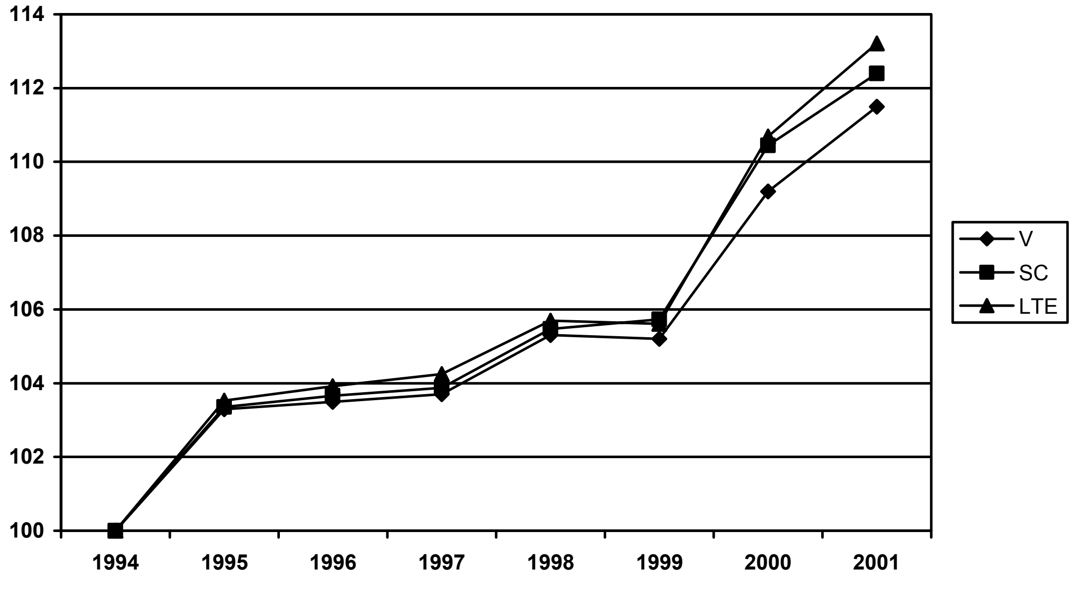
[3] A minor complications in applying this to the UK data is that the lifetime earnings estimated above refer to completed qualifications so that some pupils may attend education up to some age but not qualify. An adjustment was therefore incorporated to impute the GCSE and A-level earnings only to those pupils who gained at least one GCSE or A-level. This adjustment was minor as in the current period 98% of pupils who attend school up to age 16 or attempt A-levels achieve a pass in at least one subject.

outcome measure showing the highest growth. This outcome measure increases the annual average growth rate to 1.77 as against 1.55 using the crude volume measure.

Thus an outcome measure based on the valuation by the market suggests a small, but not insignificant, upward adjustment to the volume measure.

Note these life-time earnings calculations do not vary with time. If the improvements in test scores in Table 2, are real changes in the effectiveness of pupils then wages relative to a base category should change across time. One possibility would be to examine age cohort effects, for example considering the wages received by individuals aged t in one year relative to those aged t+1 in that year. The next subsection discusses some crude estimates based on a cohort analysis.

Figure 2 Output and outcome measures for UK Education, 1994-2001



Notes:

- *V= Volume measure, SC = measure based on test scores (weight 1.10) and LTE = measure based on lifetime earnings.*

5.2 *Cohort Analysis Example*

As an example of the methods of implementing such a cohort analysis of education outputs, and the problems associated with this, we consider the following example for the UK. The data come from the Labour Force Survey 1997-99. Consider two cohorts, one leaving school in 1995, and one leaving in 1994. We could compare the affect of their GCSE by comparing the wages they earn. Table 10 provides such a comparison. The figures for 'mean wage' is the mean net weekly wage of full time workers of a particular age taken from the summer quarter of the LFS. The summer quarter was chosen because all but one twelfth of the age group would have graduated in the same year[4]. The 'raw difference' is the proportional difference between the mean wage of that cohort and that of the cohort a year younger, i.e.

$$r_{18} = \frac{w_{18} - w_{17}}{w_{18} + w_{17}}$$

where r_{18} is the raw difference for the cohort aged 18, w_{18} is the mean wage for that cohort and w_{17} is that for the cohort aged 17.

The wage net of experience is the mean wage adjusted for the cohort's potential experience, i.e. their current age minus the age they left school (in this case, by the end of the academic year they would be 16), using the estimated coefficients from Table 4. The adjusted difference is the proportional difference between the mean adjusted wage of that cohort and that of the cohort a year younger.

We can see from Table 10 that the comparison depends on the date in which one uses to calculate the difference. If we had done so in 1996, we would say that those who took a GCSE in 1994 earned 21.3% more than those who graduated in 1995. However if we compared cohorts using 1997 data, we would say this difference was 15.6%. If we chose

1998, the figure would be -0.002%! There are a number of reasons for these disparities. First, the cell sizes are particularly small. It may well be that other factors that influence wages are not constant across the samples. Second, we may be mistaken to take the potential experience effect from an equation estimated on one year's data (for summer 1996 to spring 1997). Given these reservations we do not adjust our baseline figures to take account of cohort affects. Future extensions of this work will examine the use of panel regressions as a tool to estimate age cohort effects.

Table 10 Cohort Analysis of GCSEs on Wages

Age	17		18		19		20		21	
	Wage	*N*	Wage	*N*	Wage	*N*	Wage	*N*	Wage	*N*
1996										
Mean wage	*70.97*	*38*	**98.11**	**38**	121.31	29	122.60	30	142.89	36
Raw difference			**0.3209**		0.2115		0.0106		0.1528	
Net of experience	*63*		**78.00**		87.14		80.23		85.87	
Adjusted difference			**0.2128**		0.1107		-0.083		0.0679	
1997										
Mean wage	88.88	40	*100.92*	*39*	**130.63**	**27**	122.93	28	147.91	34
Raw difference			0.127		**0.2566**		-0.061		0.1845	
Net of experience	78.89		*80.24*		**93.84**		80.45		88.89	
Adjusted difference			0.017		**0.1562**		-0.154		0.0997	
1998										
Mean wage	95.49	76	114.72	89	*135.20*	*71*	**148.07**	**57**	150.21	47
Raw difference			0.183		0.1639		**0.0909**		0.0144	
Net of experience	84.76		91.21		*97.12*		**96.90**		90.27	
Adjusted difference			0.0733		0.0627		**-0.002**		-0.071	

5.3 Changing outcomes: Contributions of Education to raising economic growth

We now turn to a consideration of changes in the effectiveness of education across time. We have already dealt with the use of test score results and argued that without information on the extent to which reaching some threshold level raises effectiveness we

[4] A more thorough analysis would use the month of birth and the month interviewed to obtain the exact population in the particular cohort, but this example is for illustrative purposes only.

cannot use this information in a transparent way. The remainder of this section therefore discusses an alternative adjustment based on the impact of education on economic growth.

Over time average earnings have risen due to increases in labour productivity. In the decade 1990 to 2000 output per hour worked grew at an annual average rate of about 2.1% in the UK and 1.5% in the US. A total outcome measure would add these increases in real earnings to the growth in volumes at each point in time. But only a part of this increase is attributable directly to the education sector. The standard growth accounting method, most commonly associated with Dale Jorgenson and collaborators (e.g. as set out in Jorgenson et al. (1987)) divides changes in labour productivity into changes in physical capital intensity, labour quality and underlying residual productivity, most commonly termed total factor productivity (TFP). In this framework the education sector's impact is in raising labour quality through increasing the skills of the workforce. Recent estimates by O'Mahony, Robinson and Vecchi (2003) suggest the contribution of skills to labour productivity growth from 1990 to 2000 averaged 0.50 percentage points in the UK and 0.20 percentage points in the US. Thus just under a quarter of UK labour productivity growth in that period was due to increases in labour force skills whereas it only accounted for about 13% of the US labour productivity.

The final outcome figures in both countries are adjusted to take account of this impact on growth. In this first attempt we use this growth accounting estimate to increase lifetime earnings by 0.5% per annum in the UK and 0.2% in the US. Thus we are modifying the outcome flow measure in a manner equivalent to that set out in equation (9) in section 2 above. This is a crude measure in that it we assume the impact is constant throughout the decade and a more sophisticated approach will be attempted at a later date. It also assumes the impact is equal across education groups. In fact there is ample evidence of a skill bias in technological change across time and that this may be linked to adoption of information technology (e.g. see the survey in Chenells and van Reenen, 1999). Thus wages of higher skilled workers relative to the unskilled, in particular those with degrees, have tended to increase over time. If changes in technology are exogenous then we would not wish to attribute these differential earnings impacts to the education sector. On the other hand if the education sector has responded to this demand by changing the subjects taught to students to reflect increased demand for computer related skills then there is an argument for including an adjustment that varies by education level. But in practice estimating these impacts would be very difficult and so are not pursued here. Finally the growth accounting method is likely to underestimate the impact of skills on growth, if there are complementarities between physical capital and skills or if TFP growth is affected by increases in human capital through external effects or spillovers. However it is difficult to quantify these aspects and so are not attempted here.

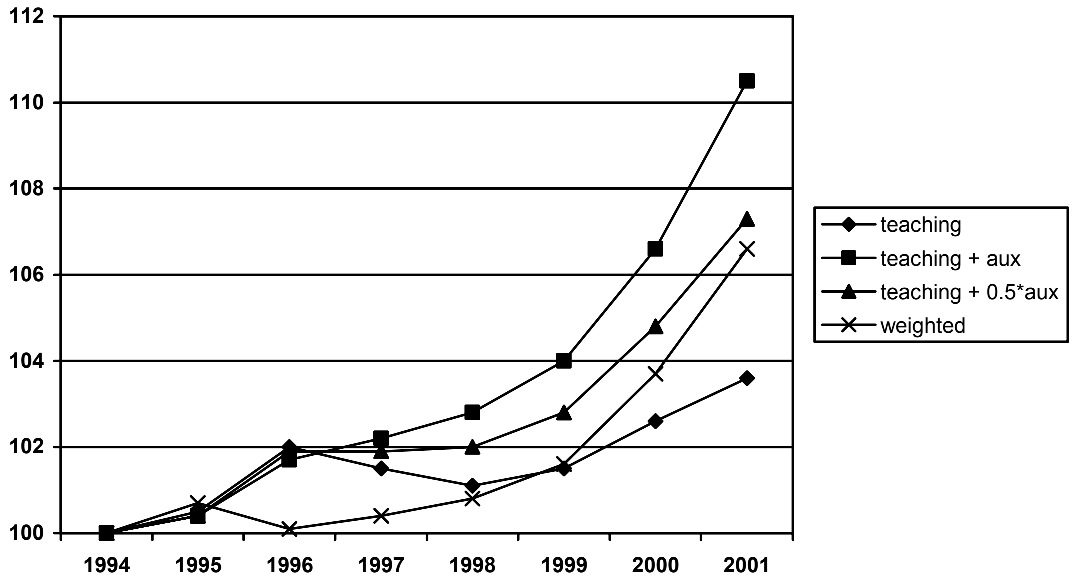
5.4 Inputs and Labour Productivity

5.4.1 Inputs

The simplest measure to start with is number of workers. The estimates will vary however depending on whether only qualified teachers are included or if we also include auxiliary staff including teaching assistants since the latter have been growing rapidly in the UK in recent years. Hence total teaching staff have been growing by 0.5% per annum from 1994 to 2001 whereas teaching auxiliaries have been rising by nearly 6% per annum on average over this period, with the growth concentrated in the final few years. An alternative measure, and the one in the spirit of the method employed to measure output, is to weight each type of staff by their wages rates relative to a base category (primary teachers).

Figure 1 plots the growth in total teaching staff, teaching staff including auxiliaries, including the latter but with a weight equal 0.5 and the wage weighted alternative. Wage weighting leads to a larger rise in labour input than just using teachers alone but to lower increases than either option including auxiliaries (the two dashed lines). In fact the latter illustrates the sensitivity of the results to the assumptions employed for auxiliary staff.

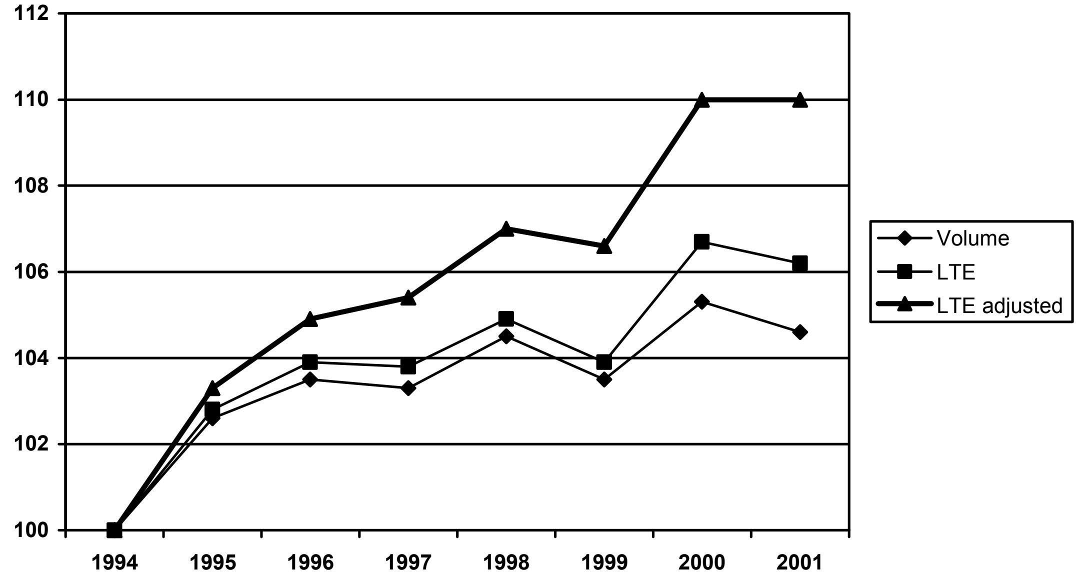
Figure 3 Labour Input in Education, UK



5.4.2 Productivity Measures

In terms of labour productivity, combining the preferred wage weighted measure of labour input, with the output measures suggests an annual average growth rate from 1994 and 2001 of between 0.64% if the volume measure is employed, 0.86% if the earnings outcome flow measured is employed as an alternative and 1.36% if the latter is adjusted for the impact on economic growth. To put this in context, economy-wide growth in UK labour productivity over this time period grew, on average, by 1.44% per annum. Thus labour productivity growth in the UK education sector was below the economy-wide measure, but only marginally so using the final measure above. However the economy wide measure does not adjust for labour quality so is not directly comparable. It is useful therefore to compare the same sector across countries. We now turn to a discussion of US results.

Figure 4 UK Labour productivity



6 US Results and international comparisons

6.1 US Results

Data for the US on school and college enrollment were downloaded from the National Center for Education Statistics web-site. Raw numbers were adjusted to full time equivalents to be consistent with the UK figures. Staff numbers and salaries came from the same source. Data underlying the lifetime earnings calculations came from the Current Population Survey.

6.1.1 *Outputs and outcomes*

In terms of international comparisons, the lack of national tests in the US means that it is not possible to construct a test score weighted index for that country. International test score results tend to be very limited, often confined to a certain age group, e.g. 15-16 year olds, and have been subject to definitional changes over time. Therefore for the US we only consider the volume and earnings outcomes measures.

Table 8 shows the discounted lifetime earnings results for the US. In comparison with the lowest category, the wage premiums are much greater in the US than those for the UK shown in Table 7 above. Proportionally, the greater US premiums are highest at the top end of the education distribution.

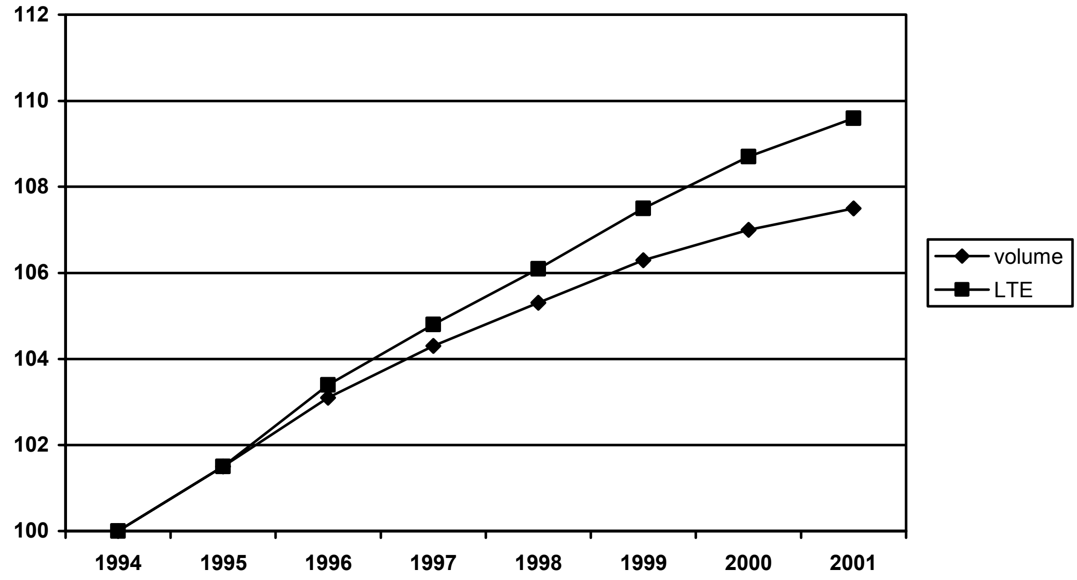
Table 11 Discounted Lifetime earnings, US.

	Discounted lifetime earnings*	
	Total	Relative to no qualifications
post graduate	$488,224	1.909
under graduate	$582,830	2.279
Associate Academic	$474,289	1.854
Associate vocational	$458,517	1.793
12th grade- diploma	$427,160	1.670
12th grade – no diploma	$323,940	1.266
11th grade	$283,337	1.108
No Qualifications < 11th grade	$255,778	1.000

* average of male and female earnings

The volume measure of output for the US shows an annual average increase of 1.03% from 1994 to 2001, which is lower than in the UK. Using the same method as for the UK to translate lifetime earnings to education outcomes leads to growth of 1.31 % per annum, a significant upward adjustment. The plot of both series in Chart shows increasing divergence across time.

Figure 5 US output growth, 1994-2001

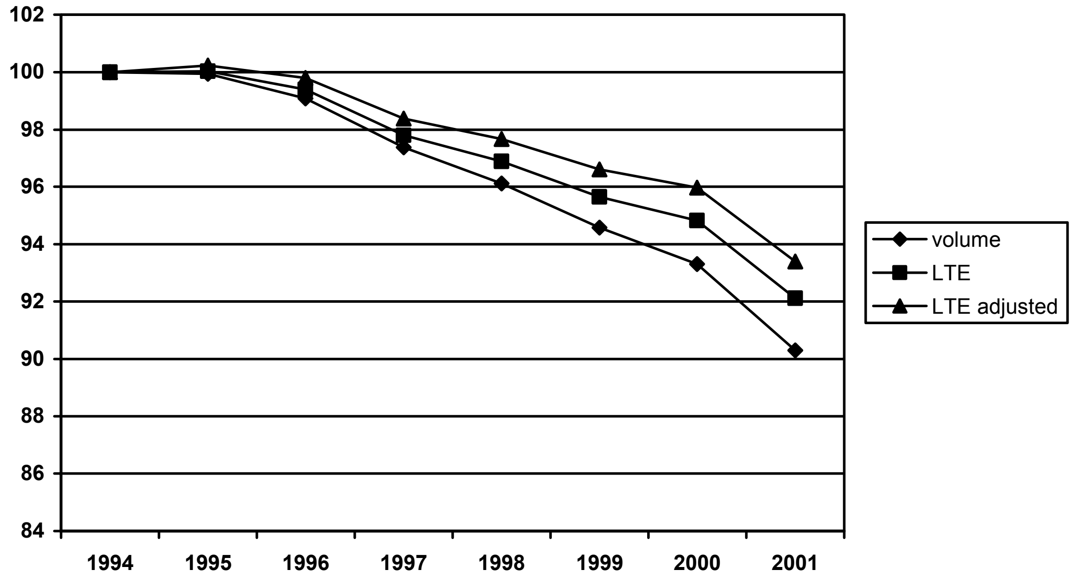


6.1.2 Labour Input and Labour productivity

As for the UK we calculated both a volume measure of labour input (number of full-time equivalent teachers and HE staff) and a wage weighted variant. In this case there was only a marginal difference between the two calculations with the volume measure increasing by 2.55% per annum between 1994 and 2001 and the wage weighted measure rising by 2.49%.

As with the UK, we also include a variant that adjusts for the impact on economic growth, raising the earnings outcome measure by 0.20% per annum. The net effect of the output and input measures is a decline in US labour productivity growth in the last half of the 1990s, as shown in Figure 6. Thus on average, the volume measures suggests annual declines of 1.47%, the LTE measure of 1.17% and the growth adjusted LTE measure by 0.97%.

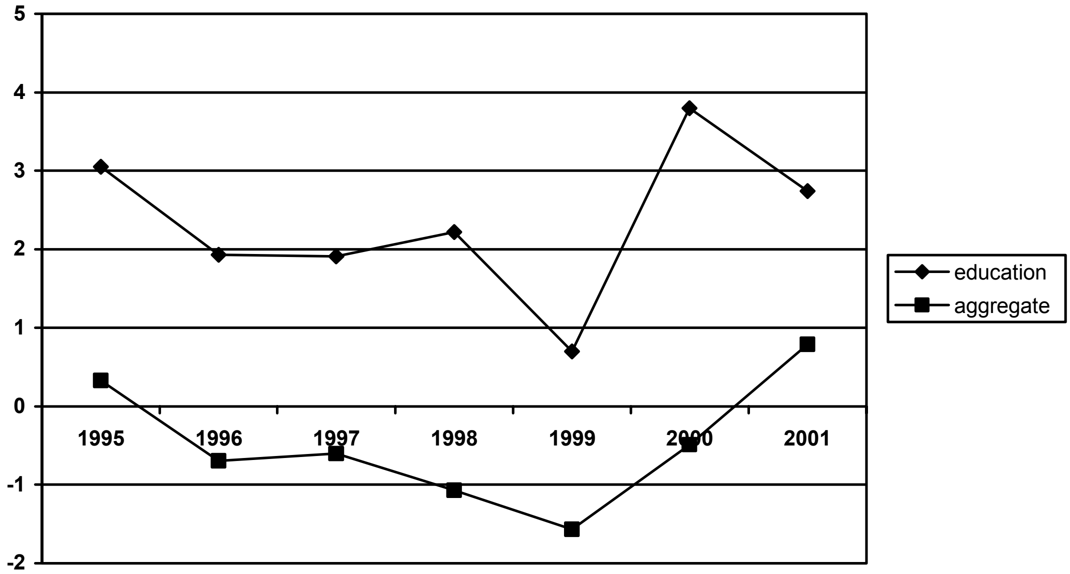
Figure 6 US Labour productivity growth, 1994 to 2001



6.2 International Comparisons

Figure 7 plots the annual percent difference between UK and US labour productivity growth rates in the education sector, using the growth adjusted outcome measure, and compares this with the difference in productivity growth rates in the aggregate economy. In Education the UK outperforms the US in all years (the differences are everywhere positive) with very high differences in the final two years. In contrast labour productivity growth in the total economy was mostly greater in the US than in the UK during these years. Hence put in this international context productivity growth in the UK education sector looks impressive.

Figure 7 Annual differences in Labour productivity growth rates, (UK-US) Education and the Aggregate Economy, 1995-2001



7 Conclusions and future extensions.

This paper is a first attempt to consider the performance of the UK education sector in an international context. The measure chosen, labour productivity growth, implies that the UK outperformed the comparable sector in the US in recent years. But arriving at this result involved a number of crude assumptions and further refinements are needed.

First it would be useful to extend the analysis back in time, e.g. to include the 1980s as well as the 1990s. In principle the data required to undertake this extension are available but there are practical problems in matching data across time. Thus in the UK case we will need to move to using an alternative source to estimate returns to education, since the Labour Force Survey only included information on wages from 1993 onwards. An alternative source is the General Household Survey which contains data back to the late 1970s but with some changes across time. The US Current Population Survey does have data on earnings back to 1976 but with a break in educational groups in the early 1990s. Nevertheless these data problems are not insurmountable and the researchers will attempt to extend the analysis in this way.

The survey data mentioned above relate to stocks of workers with various qualification levels at each point in time. A more sophisticated framework would attempt to attempt to look at flows and incremental changes to lifetime earnings as well as taking account of age cohort effects in a panel regression framework to measure changes in effectiveness of the education system through time. Finally we plan to extend the analysis to other European countries. This will draw on the survey data that have been used to estimate returns to education in Harmon et al. (2001).

References

Card., D., (1999), ‘The Causal Effect of Education on Earnings’, in O. Ashenfelter and D. Card (eds.) *Handbook of Labor Economics*, Amsterdam and New York: North Holland.

Cawley, J., Heckman, J., and Vytlacil, E., (1998), ‘Cognitive Ability and the Rising Return to Education’, NBER Working Paper Series no. 6388.

Chenells, L. and J. Van Reenen, (1999) ‘Has technology hurt less skilled workers? An econometric survey of the effects of technical change on the structure of pay and jobs’, IFS Working paper

Chevalier, A., and Walker, I., (2001), ‘United Kingdom’, in Harmon, Walker and Westergaard-Nielsen (eds.), *Education and Earnings in Europe*, Cheltenham: Edward Elgar.

Cutler, David M. and Ernst R. Berndt, eds. (2001), *Medical Care Output and Productivity*, Chicago, Chicago University Press.

Cutler, David M. and Mark McClellan (2001). Productivity Change in Health Care *American Economic Review*, 91 (2): 281-286.

Gold, Marthe, R., Joanna E. Siegel, Louise B. Russell and Milton C. Weinstein, eds. (1996), *Cost Effectiveness in Health and Medicine*. New York, Oxford University Press.

Harmon, Walker and Westergaard-Nielsen (eds.), (2001), *Education and Earnings in Europe*, Cheltenham: Edward Elgar.

Heckman, J., (1976), ‘The common structure of statistical models of truncation, sample selection, and limited dependent variables in a simple estimator for such models’, *Annals of Economic and Social Measurement*, 5, pp. 475-92.

Heckman, J., (1979), 'Sample selection bias as a specification error', *Econometrica*, 47, pp. 153-61.

Jorgenson, D.W., Gollop, F.M. and Fraumeni, B. (1987), *Productivity and US Economic Growth*, Cambridge MA, Harvard University Press.

Jorgenson, D. W., and Fraumeni, B. M., (1991), 'The Output of the Education Sector', Harvard Institute of Economic Research, Discussion Paper No. 1542.

O'Mahony, M., Robinson, C. and Vecchi, M. (2003), Information Technology, Skills and Productivity Growth: International Comparisons, mimeo, National Institute of Economic and Social Research, London.

Stevens, P.A., (2003), *Changing Patterns of UK Unemployment and Employment*, D.Phil. Thesis, University of Oxford.

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

Trostel, P., Walker, I., and Wooley, O., (2002), 'Estimates of Economic Return to Schooling in 28 Countries', *Labour Economics*, 9.1, pp. 1-16.

8 Appendices

Appendix 1 Additional earnings estimation for UK

Table 12 Additional Earnings Equations, UK

(Summer 1998 to Spring 1999)
Using Heckman Selection Method

	Men		Women	
	Earnings equation	*Selection equation*	*Earnings equation*	*Selection equation*
potexp	0.11394*** (0.00251)		0.08356*** (0.00302)	
potexp^2	-0.00426*** (0.00015)		-0.00437*** (0.00017)	
potexp^3	0.00005*** (0.00000)		0.00007*** (0.00000)	
Health problem	0.08150*** (0.01320)	-0.59787*** (0.01678)	0.36069*** (0.01685)	-0.67897*** (0.01590)
Black	-0.00446 (0.03283)	-0.33129*** (0.04543)	0.27180*** (0.03836)	-0.34403*** (0.03960)
Indian	0.03942 (0.02932)	-0.37124*** (0.04074)	0.32935*** (0.04005)	-0.43160*** (0.04014)
Pakistani/ Bangladeshi	-0.08412** (0.03820)	-0.54455*** (0.04907)	0.65663*** (0.06564)	-1.04966*** (0.05800)
Other Asian	0.05384 (0.05205)	-0.52676*** (0.06936)	0.49062*** (0.06116)	-0.55658*** (0.06046)
Mixed	0.02329 (0.06056)	-0.25835*** (0.08298)	0.42148*** (0.07417)	-0.46754*** (0.07441)
Other	0.19063** (0.08369)	-0.63735*** (0.10868)	0.57953*** (0.10431)	-0.61977*** (0.10280)
HE_PG	0.59860*** (0.02061)	0.58747*** (0.03134)	0.71400*** (0.02916)	0.55253*** (0.03226)
HE_UG	0.51835*** (0.01494)	0.46083*** (0.02105)	0.55186*** (0.01844)	0.50678*** (0.01898)
FE	0.13698*** (0.01313)	0.42453*** (0.01788)	0.09353*** (0.01562)	0.48974*** (0.01574)
TRADEAPP	0.15540*** (0.01499)	0.18765*** (0.02070)	0.09352*** (0.02943)	0.18148*** (0.03039)
ALEVEL	0.29164*** (0.01758)	0.30376*** (0.02452)	0.29188*** (0.02102)	0.35932*** (0.02158)
GCSE	0.01328 (0.01413)	0.39084*** (0.01948)	-0.02545* (0.01547)	0.41592*** (0.01562)
DKQUAL	0.73316*** (0.03391)	-1.26907*** (0.03554)	1.20224*** (0.04407)	-1.11242*** (0.03722)
age		0.17533*** (0.01318)		0.02453** (0.01232)
age^2		-0.00405*** (0.00037)		0.00062* (0.00035)
age^3		0.00003*** (0.00000)		-0.00001*** (0.00000)
married		0.22931*** (0.01068)		-0.08884*** (0.00824)
Constant	8.77365*** (0.01771)	-2.44973*** (0.14627)	8.83269*** (0.02083)	-1.01683*** (0.13741)
ρ	-0.87		-0.94	
σ	0.70		0.97	
λ	-0.61	(0.01)	-0.92	(0.01)
χ^2	2183.56		3394.78	
$p(\chi^2)$	0.00		0.00	
Observations	57929		61144	
censored	28175		30070	

- *Standard errors in parentheses*
- ** significant at 10%; ** significant at 5%; *** significant at 1%*
- χ^2 *= Likelihood ratio test of* $\rho = 0$

Table 13 Additional Activity equations, UK

(Summer 1998 to Spring 1999)
Multinomial logit (omitted category = in employment)

	Men		Women	
	Unemp	*Inactivity*	*Unemp*	*Inactivity*
age	-0.30952*** (0.04446)	-0.82446*** (0.03569)	-0.17363*** (0.05476)	0.20635*** (0.02525)
age^2	0.00649*** (0.00130)	0.01779*** (0.00102)	0.00352** (0.00162)	-0.00851*** (0.00071)
age^3	-0.00004*** (0.00001)	-0.00011*** (0.00001)	-0.00003* (0.00001)	0.00009*** (0.00001)
Married	-1.20281*** (0.04887)	-0.76913*** (0.03954)	-0.79368*** (0.05457)	0.04615* (0.02393)
Health problem	1.16001*** (0.05190)	2.80280*** (0.03545)	0.95970*** (0.06436)	1.71260*** (0.02747)
Black	0.85164*** (0.12108)	1.14144*** (0.10569)	0.91868*** (0.12024)	0.54300*** (0.07515)
Indian	0.40098*** (0.13604)	0.70546*** (0.10319)	0.59995*** (0.15017)	0.67699*** (0.07096)
Pakistani/ Bangladeshi	1.05233*** (0.12522)	1.37685*** (0.10227)	1.48856*** (0.17908)	2.09429*** (0.09403)
Other Asian	0.52560** (0.23970)	1.69398*** (0.14896)	0.57014** (0.24846)	1.15691*** (0.10331)
Mixed	0.57591*** (0.22164)	0.64708*** (0.19384)	1.21377*** (0.19740)	0.67713*** (0.13511)
Other	1.61750*** (0.27527)	2.24742*** (0.21808)	1.10123*** (0.34217)	1.32847*** (0.17541)
HE_PG	-1.68234*** (0.15590)	-1.28598*** (0.11526)	-0.75494*** (0.14403)	-1.80776*** (0.08229)
HE_UG	-1.26554*** (0.07603)	-1.06772*** (0.06345)	-0.99350*** (0.08882)	-1.54062*** (0.04076)
FE	-0.88444*** (0.05329)	-1.00137*** (0.04465)	-0.62574*** (0.06544)	-1.22945*** (0.02996)
TRADEAPP	-0.94136*** (0.07156)	-0.95447*** (0.05565)	-0.69314*** (0.15022)	-0.72706*** (0.05712)
ALEVEL	-1.05940*** (0.08221)	-0.05312 (0.05622)	-0.93172*** (0.09566)	-0.83172*** (0.04052)
GCSE	-0.99302*** (0.05778)	-0.93228*** (0.04839)	-0.66771*** (0.06442)	-0.91202*** (0.02887)
DKQUAL	-0.87256*** (0.15381)	-0.70100*** (0.12476)	-0.59792*** (0.19614)	-0.91949*** (0.09166)
Constant	2.96936*** (0.46881)	9.43879*** (0.37928)	0.70079 (0.57032)	-1.36583*** (0.27672)
Observations	55366	55366	59387	59387

- *Standard errors in parentheses*
- *significant at 10%; ** significant at 5%; *** significant at 1%*