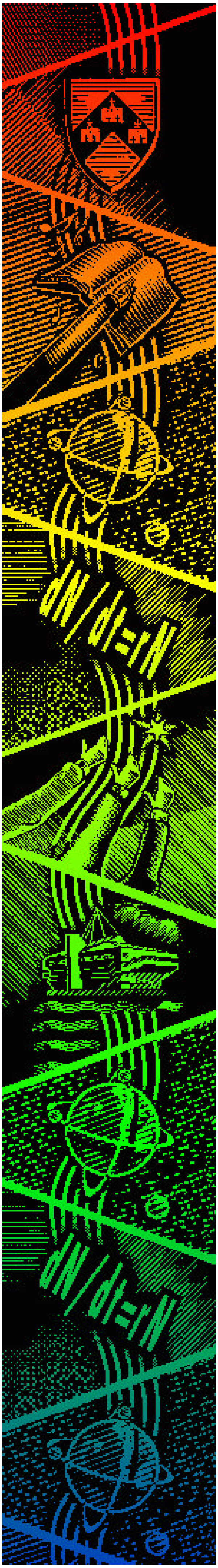

THE UNIVERSITY of York

Discussion Papers in Economics

No. 2000/40

Performance Management and Performance Measurement in the Education Sector

by

David Mayston

Department of Economics and Related Studies
University of York
Heslington
York, YO10 5DD

PERFORMANCE MANAGEMENT

AND

PERFORMANCE MEASUREMENT

IN THE EDUCATION SECTOR*

DAVID J. MAYSTON**

University of York

* This is a revised version of the paper presented by the author to the OR42 Annual Conference of the Operational Research Society, Swansea, 14th September 2000.

** The author is Professor of Public Sector Economics, Finance and Accountancy, and Director of the Centre for Performance Evaluation and Resource Management, at the University of York.

ABSTRACT

The paper examines several outstanding issues on the interface between the measurement of performance in primary and secondary education and the management of improved performance in this nationally important sector. These issues relate to the clarification of the objectives of the education system, the impact of performance reward systems, such as Performance Related Pay, the role of resources in influencing educational outcomes, the reliability of existing methods of assessing educational performance, such as Data Envelopment Analysis and multivariate regression, and the need for an improved national comparative database if progress is to be made in several of these directions.

Key words: performance management, performance measurement, education, data envelopment analysis, quality control, knowledge management.

Address for correspondence: Department of Economics and Related Studies, University of York, York, Y010 5DD, UK.

E-mail: dm3@york.ac.uk

Introduction

Education provides an important area of application for systems of performance management and for techniques of performance measurement. Education is currently an area with a high national priority, in the form of a high weight given by politicians on the raising of educational standards and performance, particularly in primary and secondary education. However, the careful development of performance management systems and performance measurement techniques poses a number of detailed analytical problems that take on particular significance in the education sector, and which merit further examination. At the same time, interesting issues are raised about how best to increase and manage knowledge and information in the process of raising educational performance.

Clarification of objectives

A key potential role for non-profit performance indicators in public services, such as education, is that of clarification of the objectives of each service[1]. The need for such clarification becomes even more relevant if performance measures are deployed within performance reward systems, such as Performance Related Pay (PRP), or within performance management systems that make public judgements on individual schools and teachers, such as in publicly available OFSTED[2] school inspection reports in the UK. These systems can provide powerful incentives for individual schools and teachers to seek to maximise their reported performance according to the measurement framework that is imposed upon them by the performance reward or management system. However, there are a number of systematic ways in which this may lead to 'sub-optimisation' of the educational outcomes compared to wider social goals.

The first is a neglect within the reported performance measures of one or more of the dimensions of educational outcomes that are actually valued by society at large. These dimensions might include some measure of the extent of pupils' fulfilment or satisfaction that pupils themselves derive from the large proportion of their lives they spend in education. There may then exist a *non-monotonic* underlying relationship between pupils' satisfaction and reported examination

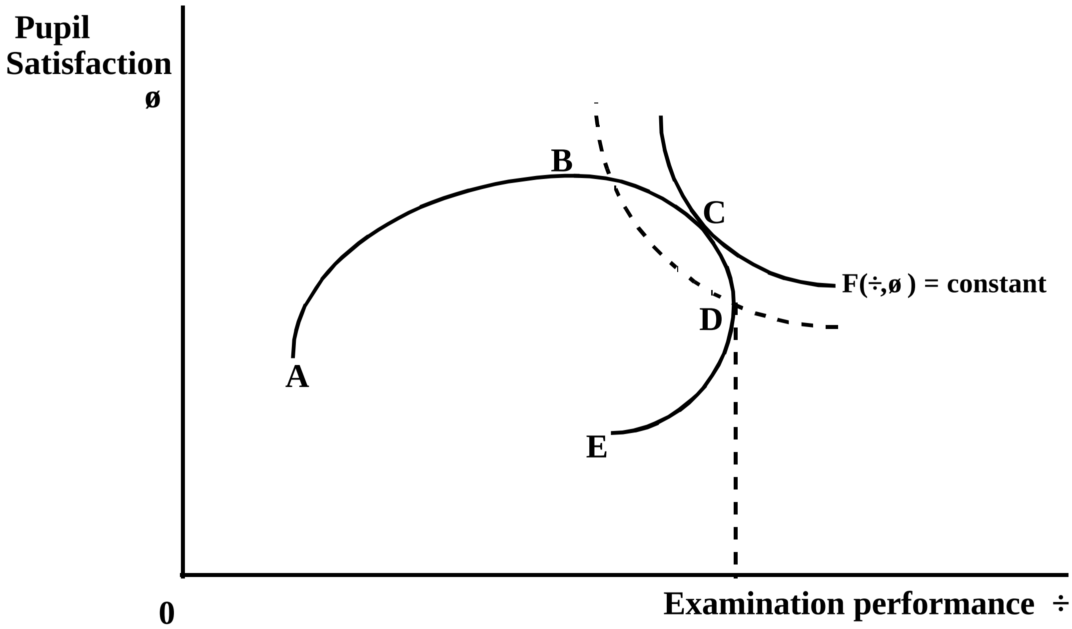

Figure 1

results, as in Figure 1. Over some range, such as from A to B in Figure 1, pupils' satisfaction may increase when they are the subject of more education aimed at improving their examination results. However, after some point the additional stress and pressures that they suffer from further increasing their examination performance reduces their satisfaction and sense of fulfilment from the educational process itself, as between points B and D in Figure 1. Even greater pressure may impair also their examination performance, as between points D and E. An overall evaluation function F(÷, ø) that pays attention both to the level of their examination results, ÷, and to the level of pupil satisfaction, ø, may achieve its maximum point subject to the underlying relationship between ÷ and ø at a point such as C in Figure 1. Maximising examination results at point D will lead to a lower overall value of F than at C, by trading-off in a sub-optimal way (according to the evaluation function F) reduced pupil satisfaction for higher examination results. In contrast, a performance management system that places value only on examination results and deploys a frontier technique, such as Data Envelopment Analysis (DEA)[3, 4], will identify a school at point C as inefficient compared to a school at point D within the subspace defined by all feasible values of examination results, ÷. A wider issue raised for performance management systems is then how far improved reported performance is genuinely the result of increased organisational efficiency inside the feasible frontier of achievement, and how far it is instead due to moving along a frontier in a larger dimensional space that involves sacrificing less easily measurable outcomes, such as pupil satisfaction, self-fulfilment or enjoyment, which may still nevertheless be valued by some important stakeholders in the production process. Whilst this issue arises even when there is a monotonic negative relationship between the two outputs, a change in the direction of the relationship can undermine any initial assumption that there is no conflict between the two outputs that is made when they are initially positive related under conditions of low stress.

The relative evaluation of pupil satisfaction or fulfilment compared to examination results also raises complex issues concerning the relative importance, and size, of the short-term and long-term benefits that are derived from the educational process. Whilst this in part involves questions of the discount rate that should be applied in present value calculations for these future benefits compared to current benefits, underlying issues are also involved concerning the extent of the future benefits which will be obtained by different groups of pupils from improved examination results. The relative valuation of improved examination results and future national achievements in sport may also arise if increased pressures on pupils and teachers to boost examination results diminishes the time devoted to active participation in sporting activities.

Achieving an overall optimal level of performance for the education system within finite total resources for the educational system is itself likely to involve *trade-offs* between the resources devoted to different groups of pupils, and the resultant benefits which these different groups obtain from the additional educational resourcing. Many existing statements of objectives for the educational system are formulated in terms which recognise no such trade-offs. Thus the UK’s Department for Education and Employment (DfEE)[5] states its aim as to be “to give everyone the chance, through education, training and work, to realise their full potential, and thus build an inclusive and fair society and a competitive economy”. Under certain conditions, frontier performance measurement techniques, such as DEA, can be used to assist in the process of making all groups within the educational system better off, by identifying the scope for *Pareto improvements*, i.e. movements to feasible vectors of educational outcomes for each relevant pupil group on the DEA efficiency frontier that *dominate* the existing organisational achievements for each of these pupil groups. They can also help to identify the nature of the trade-offs which are involved on the production side of the educational process along this efficiency frontier between the educational outcomes for these different groups. In doing so, they can help to make policy judgements on how to allocate resources between these different groups more explicit and better informed. However, the identification of both the extent of the possible Pareto improvements and the shape of the efficiency frontier is contingent upon the available information being organised in a suitable *disaggregated* form that enables the achievements of different pupil groups themselves to be identified.

Without such disaggregation, reliance upon current standard measures of school achievements, such as the proportion of pupils achieving five or more GCSE grades A* - C, is likely to involve a number of more questionable *implicit* policy trade-offs between the achievements of different pupil groups. Linking these standard performance measures to performance reward systems, such as PRP, is in turn likely to cause these implicit, and possibly ill-considered, trade-offs to drive educational outcomes. Providing an incentive for a school or teacher to maximise the proportion of pupils achieving five or more GCSE grades A* - C itself encourages the development of an internal performance management system aimed at 'managing the margins'[6] of pupils who are on the borderline of achieving just five GCSE grades A* - C. Directing additional resources and attention to them, and away from those who are unlikely to be raised to this level and away from those who are well above this borderline of achievement, is the performance management strategy most geared to the goal at hand, albeit at a cost to wider educational objectives. That this ill-designed performance indicator, which is prominent also in the targets set for the education system as a whole[7, 8], is increasingly dominating many educational policies within schools is confirmed by a recent study by Gillborn and Youdell[9].

Avoiding such implicit and questionable trade-offs, however, requires more explicit attention to be given to sensitive issues concerning the attitude which the educational system should take to *inequality* of educational achievements. Once recognition is given to overall educational resource constraints, the question arises as to whether the educational system should devote at the margin more resources to those at the lower end of the educational spectrum of examination results, or to those in the centre, or to the pursuit of excellence by those at the upper end of the spectrum.

Educational performance management systems that operate under the assumption that there are still large Pareto improvements to be had within the educational system can to some extent avoid this question. However, the greater the success that is achieved by the performance management system in realising these Pareto improvements, the greater is the need to recognise the policy issues that arise once schools are on the efficiency frontier and optimal resourcing choices have to be made along this frontier of educational achievement.

There are, fortunately, economic reasons for believing that any strongly unequal system of educational resource allocation is itself likely in the long term to be *Pareto inefficient* for all pupil groups, once account is taken of the long-term *pecuniary externalities* that may exist between the different groups. Neglecting educational resources and achievement at the lower end of the spectrum is likely to lead to large numbers of poorly educated, unqualified and disaffected individuals who are in a weak position in the labour market, and who may impose high long-term costs on the rest of the economy through increased social security payments, increased crime rates[10] and other social problems. Neglecting educational resources and achievement at the upper end of the spectrum is likely to undermine overall national economic competitiveness and the tax base from whom other members of the community can benefit. Even a 'maxi-min' policy of 'maximising the welfare of the worst-off individual'[11], involving an extreme aversion to inequality, can imply relatively low optimal marginal income tax rates on the highest income earners[12] in order not to deter them from earning more and contributing more tax revenue to finance greater benefits for the worst-off individual. Similar arguments are likely to limit the extent to which it is in the interest of all current pupil groups to under-resource high educational achievers.

A further source of implicit educational weights is that of the equal weighting in reported School Performance Tables[13] of examination results in different subjects, such as business studies, economics, physics and mathematics. Such equal weighting in itself implies an *indifference* in the valuation of educational outcomes under performance reward systems that respond to aggregate point scores irrespective of the subject involved, either for the school or for individual pupils seeking admission to higher education on the basis of their A-level grades. However, such equal weighting is itself likely to encourage a switching of pupils out of subjects which are relatively more demanding in their technical level and of greater perceived difficulty for achieving target grade levels, despite the possibly adverse impact this switch may have upon many traditional areas of UK excellence. Thus an important feature of educational change in recent years has been substantial relative reductions in the numbers of pupils taking A-level economics (which has dropped from 6.6% of A-level entries in 1990 to 2.6% in 1999[14]), physics (which has dropped from 6.6% of A-level entries in 1990 to 4.6% in 1999), and mathematics (which has dropped from 11.6% in 1990 to 8.7% in 1999), together with increases in pupil numbers taking A-levels in less technical and less mathematical subjects, such as business studies (which has risen from 1.8% in 1990 to 4.6% in 1999[14]). Similarly at GCSE, the school league table goal of maximising GCSE grades A* - C may be more easily achieved in mathematics itself by entering students for the less technically demanding 'restricted-grade' GCSE examination. Whilst a grade B is the maximum possible mark in this examination, a grade C is usually judged as easier to attain for marginal students than in the unrestricted-grade GCSE examination that has a higher technical content.

Placing more explicit differential weights in performance evaluation upon different subjects and upon different pupil groups, however, raises issues of the scope for linking performance management systems to systems-wide *strategic management* considerations of where the educational system as a whole should be headed, particularly in its interface with the future labour market needs of different possible areas of economic specialisation for the economy as a whole. Prominent authors[15,16] on strategies for boosting competitive advantage emphasise the importance of *clusters* of *specialised* inputs, such as a ready availability of well-qualified students from specialised educational institutions able to recruit well-qualified staff, for reinforcing and maintaining the competitive advantage of firms and the economy as a whole in particular specialised directions. International competition is likely to reward those countries which are well-organised in identifying and boosting the directions in which they can command an international competitive advantage, and penalise those countries which are not well-organised. If this competitive advantage involves a high level of technical expertise which requires strong mathematical proficiency and greater linkages between the educational system and vocational and professional training[17], then long-run issues of the relative priorities of the educational system are indeed raised which must help to drive its performance management system.

However, specialisation may also bring with it additional risks, such as increased exposure to relative price changes and exchange rate movements that may reduce the future economic value of particular educational specialisms. Despite several discernible systematic changes in employment patterns in recent years, such as a reduction in middle management opportunities through 'delayering' in traditional industries such as banking, and an increased globalisation of economic markets, considerable uncertainty and volatility remain important features of the world economy. Such uncertainty undermines any attempt at linking the strategic management of educational objectives to a mechanistic view of manpower planning based upon a definitive prediction of future labour market needs. Educational objectives instead need to balance the development of broadly-based aptitudes, including numeracy, literacy and computer skills, that have generic applicability to a wide range of labour markets and can reinforce *labour market flexibility and adaptability*, with more specialised teaching that can support the subsequent development of economically valuable specialist skills. An examination system that adequately tests both generic and specialised skill and knowledge acquisition is then a critical input into the educational performance management system. Rather than relying upon aggregate examination results regardless of subject, there is a need to monitor performance in each relevant direction, and arguably to restrict the weights which are used in an evaluation technique such as DEA to non-trivial values on critical skills[18].

Pursuit of the objective of maximisation of aggregate examination results is, however, encouraged the performance reward system that is implicit in the funding mechanism for individual schools in England and Wales. This requires at least 80 per cent of the school's income from its Local Education Authority (LEA) to be based upon the school's *(age-weighted) pupil numbers*[19]. Schools that are more successful in this objective will tend to encourage more parents who have access to School Performance Tables to enrol their pupils, and thereby boost the school's income under this funding system. If there are *increasing returns to scale and scope* in the production of school examination success, the boost in the school's total resources will enable it to invest in more specialised staff and facilities, as well as to attract higher quality staff, and to offer a broader range of specialised teaching, further boosting its educational performance. In contrast, those schools with low levels of examination performance will suffer falling income and pupil numbers, high average costs per pupil (once fixed costs and loss of scale economies are taken into account), and a reduction in the quantity and quality of specialised staff they can attract or retain under the resultant school budgetary pressures, making their task of performance improvement more difficult. As a performance management system, it therefore has a potentially destabilising *positive feedback loop* that tends to be dysfunctional for the performance of individual schools that fall behind in the above objective, and instead diverts resources to those schools that are initially ahead in this objective. The use of unadjusted aggregate examination results within the performance reward system also tends to reward schools that are more selective in their pupil intake and examination entries in favour of more able pupils, to the likely detriment of the educational chances of less able pupils. Again, the importance of making a well-considered choice of objective and associated performance measure is underlined.

The use of more sophisticated methods of performance assessment, such as DEA, can fortunately potentially overcome many of the defects of existing league table performance measures. DEA permits the use of characteristics of the pupil intake as input variables that can be used within the overall assessment of each individual school and LEA's performance [3,20]. These characteristics can include socio-economic variables that are believed to be related to the extent of educational disadvantage the pupil's background and circumstances involve. A school with high examination results but an advantaged pupil intake will then not necessarily be judged educationally more effective by DEA than a school with the weaker examination results and a disadvantaged pupil intake. The use of pupil *post codes* linked to Census Enumeration Districts can in principle facilitate such an analysis using disaggregated pupil-level scores and input data, as well as for more aggregated school and LEA level data. An alternative approach here is the use of individual pupil *prior attainment* scores in earlier school tests to characterise the characteristics of the pupil intake. Such scores have been extensively used in *value added* assessments[21] of school performance that rely not on the multi-dimensional framework of DEA, but instead on a comparison of the actual examination scores of individual pupils with what they would predicted to have achieved given their prior attainment scores on the basis of a national sample. *Multi-level* models[22] seek to examine the quantitative magnitude of the LEA, school and pupil-level influences on the extent of individual pupils' educational value added.

The role of resources in education

A central question in educational performance assessment and performance management is the role of resources in influencing educational outcomes. The present UK Government has invested a claimed additional £19 billion[5] in education in the expectation of improved educational outcomes from the UK education system. The DfEE has entered into Public Service Agreement[7] with HM Treasury with promises of delivering quantitative targets of improved educational performance, including an increase in the proportion of those aged 16 who achieve 5 or more GCSEs grades A* - C from 45 per cent to 50 per cent, in return for this additional funding. The target of a further four percentage points increase in this performance indicator forms part of the new Public Service Agreement[8] for the latest Spending Review 2000[23] that promises to increase spending on education and training by £10 billion by 2003-4.

One of the main rationales for the introduction of the system of Local Management of Schools (LMS) in 1990 was the devolution of educational budgets down to individual schools on the presumption that local head-teachers and school governors were best placed to make the more educationally effective use of their allocated resources. Achieving educational *value for money* in their use of resources is an important performance criterion both for LEAs in their required pursuit of *Best Value*[24] and for individual schools within their OFSTED inspections[2]. If there is a link between educational resourcing and educational performance, then an overall educational performance management system should understand the nature of this link, in order to optimise how the total available resources are allocated to individual schools. This allocation may take place via explicit funding formulae, such as under those of LMS from LEAs to schools and the Standard Spending Assessment (SSA) system from central government to local authorities[19], or through a new common national funding formula for all schools.

However, Hanushek[25, 26] has persistently claimed that educational resources have no apparent impact on educational outcomes, arguing that "the research of the past quarter century into educational input-output relationships has indicated clearly that schools around the world pursue very inefficient policies"[26]. Whether or not this claim is really true is of some importance both for educational performance measurement and for the appropriate response of the educational performance management system. Hanushek's claim is based upon a belief that his and other empirical studies of the *multivariate regression* relationship between school examination performance and other variables, such as expenditure per pupil and socio-economic variables, is estimating an underlying '*educational production function*'. Frontier techniques, such as DEA and stochastic frontier analysis, would, however, distinguish between the performance of schools on the frontier from those inside an efficiency frontier, and therefore away from the relevant educational production function that describes this frontier. The fact that there may be less efficient schools inside the frontier does not negate the existence of a positive relationship between resources and educational outcomes, for a given vector of pupil intake characteristics, along the efficiency frontier and within its associated educational production function.

Linking multivariate regression relationship to simply the *supply-side* concept of an educational production function can also be shown to neglect the importance of the *interaction* between the *supply-side* consideration of the additional *marginal cost* of additional educational performance and the associated *demand-side* concept of the '*willingness to pay*' for such additional examination performance[27, 28]. Balancing these relationships through an *optimal* allocation of resources by an individual school will imply that examination performance and expenditure per pupil are determined *endogenously* at the point of intersection between the local demand curve and the marginal cost curve for examination performance for a given pupil group, as in Figure 2. We would expect this marginal cost to be positive and upward sloping over some range, reflecting an underlying *law of diminishing returns* in which additional examination performance becomes progressively more costly to achieve for a given pupil intake. We would also expect this marginal cost curve to be higher for pupils from more disadvantaged backgrounds, as in the curve MC' compared to MC in Figure 2, reflecting the greater attention and educational resourcing that they need to overcome this disadvantage.

That this marginal cost relationship (in which resources have a positive role in influencing educational outcomes) is consistent with an apparent lack of any observed cross-sectional relationship between resources and educational outcomes can be seen from Figure 2. Here we assume a demand side willingness to pay for greater educational equality, such as under educational resource allocation formulae that make offsetting adjustments to school funding to compensate for the *additional educational need* [19] of a disadvantaged pupil intake. The fact that the demand curve, D'D', is also higher when the marginal cost curve is higher for this disadvantaged group in Figure 2 leads here to an equal level of examination performance, despite higher expenditure per pupil for the disadvantaged pupils.

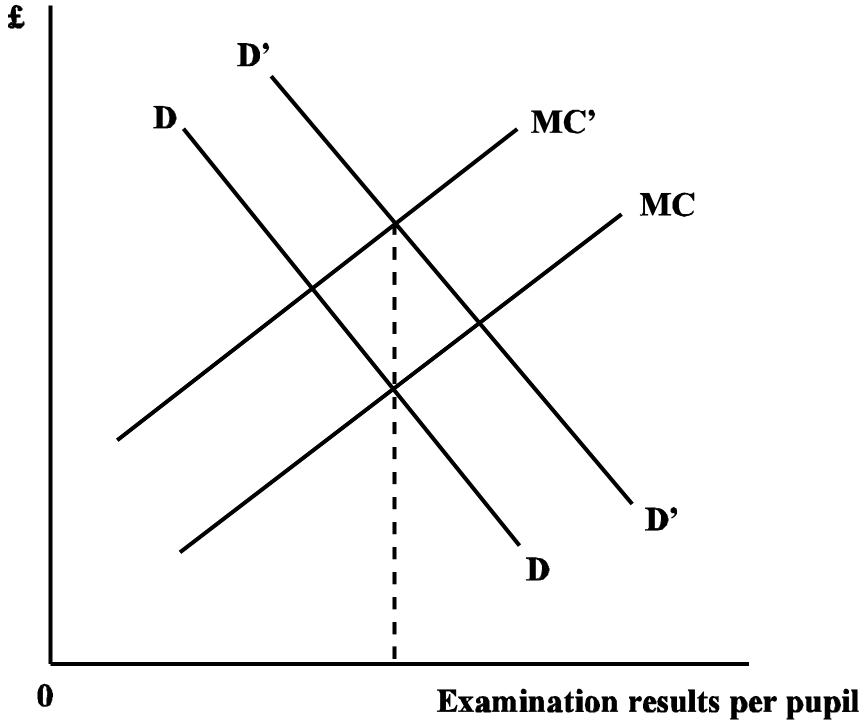

Figure 2

The demand-side inter-relationship that the resource allocation formulae may introduce between expenditure per pupil and the socio-economic variables that characterise educational advantage may imply a high degree of *multicollinearity* between expenditure per pupil and these socio-economic variables. This in turn will weaken the precision and reliability of the parameter estimates[29] of the relative contributions of expenditure per pupil and the socio-economic variables in contributing to educational performance. No great reliance can then be placed in earlier UK single-equation regression studies[30,31] which found no significant relationship between many levels of examination performance and expenditure per pupil, but which used the same socio-economic variables within the multivariate regression equations as were used within the educational resource allocation formulae to determine expenditure per pupil.

However, even aside from problems of multicollinearity, the *endogenous* nature of examination performance and expenditure per pupil under the above supply and demand interactions will bias[29] the estimated regression coefficient of the impact of expenditure per pupil on examination results that is produced by standard Ordinary Least Squares (OLS) estimation away from its true value in the underlying educational production function. These interactions will be further increased if there is a tendency for more affluent parents to migrate to the catchment areas of schools with higher published examination results and higher levels of resources per pupil. These interactions will be further increased if local house prices are pushed up in the process, so that a there is a *pre-selection mechanism* which tends to filter out more disadvantaged families from such schools. There will then tend to be a positive correlation *on the demand side* between the socio-economic background of pupils and examination results which will tend to bias upwards the estimated regression coefficient on socio-economic background in the estimated 'educational production function' under OLS away from its true value, and bias downwards that of expenditure per pupil[32]. The problem of endogeneity bias is, moreover, not restricted to OLS regression analysis, but can also reduce the reliability of frontier performance assessment techniques, such as DEA[33].

The simultaneous equations nature of the different inter-relationships which describe an *equilibrium* level of educational performance, resourcing and socio-economic intake of each school will also pose *identification* problems[29] in seeking to correctly estimate the parameters of the different equations involved. If an educational funding system does involve seeking to compensate schools for those adverse factors which raise the costs of attaining a target level of performance, then the same variables will tend to affect both the supply side and the willingness to pay side of the relationship. It may then be impossible to separately identify the quantitative parameters of the underlying supply-side educational production function, as a performance yardstick against which to judge the efficiency of individual schools.

The reliability of existing performance measurement techniques

The need and scope for assessing the reliability of existing performance measurement techniques in education arises also in the context of two main assumptions of DEA. A central implicit assumption of the constant returns to scale model of DEA is that of *homotheticity* of the estimated production structure[34]. This means that multiplying the input-output vector of an efficient school by a positive constant along a ray through the origin does not change the marginal rates of transformation (MRT) along the efficiency frontier between any pair of variables in the input-output vector. This is illustrated in Figure 3 where the MRT between any two inputs is given by the slope of the efficient isoquants. More generally, these MRT are given by the ratio of the shadow variables in the linear program (LP) formulation[35] of DEA:

$$\max\ e_0 = \sum_k s_k.y_{k0} - \sum_i u_i.z_{i0}\ \ s.t.\ \sum_h v_h.x_{hj} + \sum_i u_i.z_{ij} \geq \sum_k s_k.y_{kj}\ \ for\ j = 1,\ldots,\ n$$

$$\sum_h v_h.x_{h0} = 1;\ s_k \geq 0\ \ for\ all\ k = 1,\ldots,\ m\ ;\ v_h \geq 0\ \ for\ all\ h = 1,\ldots,ç \tag{1}$$

where y_{kj} , x_{hj}, z_{ij} are the levels of output k, controllable input h and environmental input i respectively of school j, s_k, v_h and u_i are the corresponding shadow variables, and j = 0 is the school whose technical efficiency is being estimated. The ratios of the values of s_k, v_h and u_i which solve (1) remain the same if we multiply all the y_{kj} , x_{hj}, z_{ij} by the same positive constant.

In essence, the constant returns to scale model of DEA implies a *single reference efficient isoquant* from the given sample of schools whose shape remains the same for all levels of output, subject only to a radial expansion from the origin, as in Figure 3. Under the variable returns to scale version of DEA, the same remains true, except that the rays need not pass through the origin but instead may involve a non-constant intercept with the axes[35].

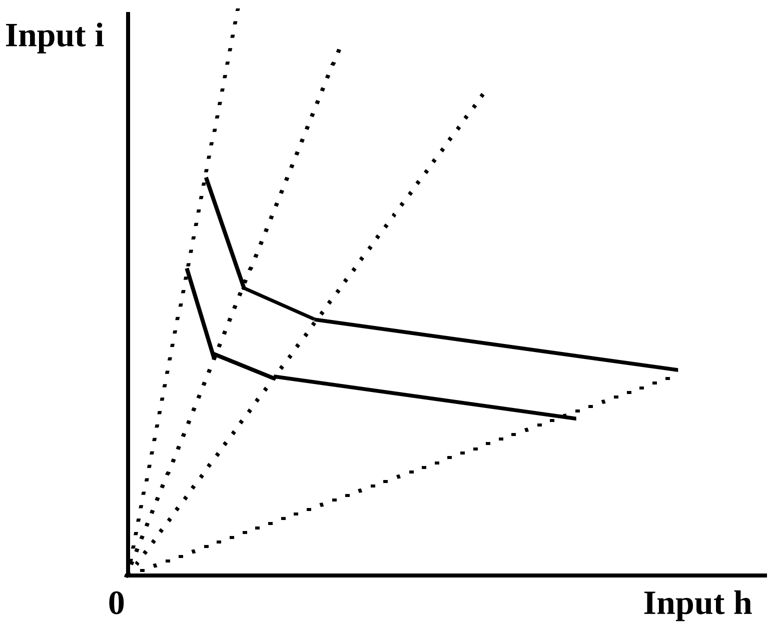

Figure 3

When we compare these implicit assumptions with the realities that may prevail in the education sector, it is quite possible that, as we expand both esources per pupil and the socio-economic parameters along a ray across the efficient isoquant map, their MRT may in fact change. A greater increase in resources per pupil may be required to compensate for a given unit reduction in the socio-economic input parameter when pupils are from very disadvantaged backgrounds than when both these variables are at higher levels along the ray. Thus in Figure 4a, the slope of the chord between schools A and B is greater than that between C and D, where schools A, B, C and D are all efficient schools and A = (8, 4, 10), B = (4, 10, 10) , C = (4, 3, 5) and D = (2, 4, 5), where the first element in the vector is expenditure per pupil, the second is the level of the socio-economic variable, and the third is examination performance.

The dual[35] of (1) is the LP:

$$\min\ \text{è}_0 \quad s.t. \sum_{j>0} w_{0j}.x_{hj} \ \leq\ \text{è}_0.x_{i0} \quad for\ \ h=1,...,\text{ç} \qquad (2)$$

$$\sum_{j>0} w_{0j}.z_{ij} \le z_{i0} \ for \ i=1,..., \ m; \quad \sum_{j>0} w_{0j}.y_{kj} \ge y_{j0} \ for \ k=1,..., \ n$$

where the $w_0 = [\ w_{0j}\ ]$ is the vector of weights applied to the input-output vectors of each school j > 0 in the sample to compute the technical efficiency coefficient of school zero. By permitting each w_{0j} to vary, DEA scales up or down the input-output vector of each school j until a convex frontier of the resultant points can be formed with which to compare the performance of any selected school 0. Scaling down points A and B in Figure 4a to points A' and B' to a comparable level of output as those of points C and D produces a set of points whose convex frontier from below for the output level 5 will not include all of the points A', B', C and D' once homotheticity does not hold in practice. Instead the convex frontier which DEA constructs in Figure 4b for the output level 5 is A' D, causing school C to be wrongly identified as having a positive slack in the socio-economic variable in the input minimisation version of DEA in (1) and (2), and to have a technical efficiency coefficient less than one in the corresponding output maximisation[35] problem.

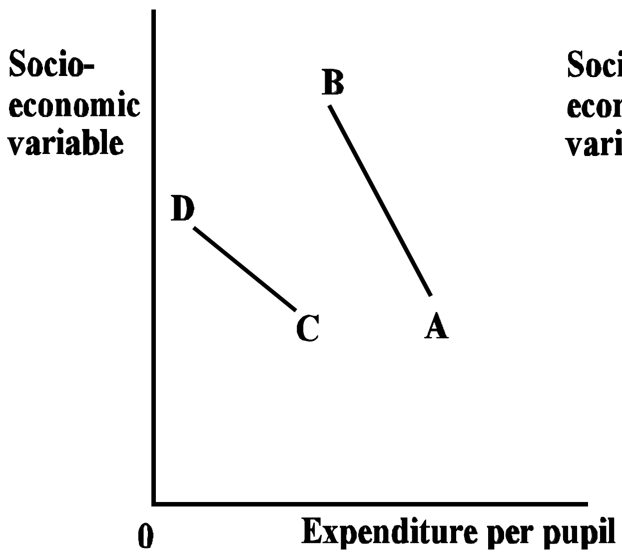

Figure 4a

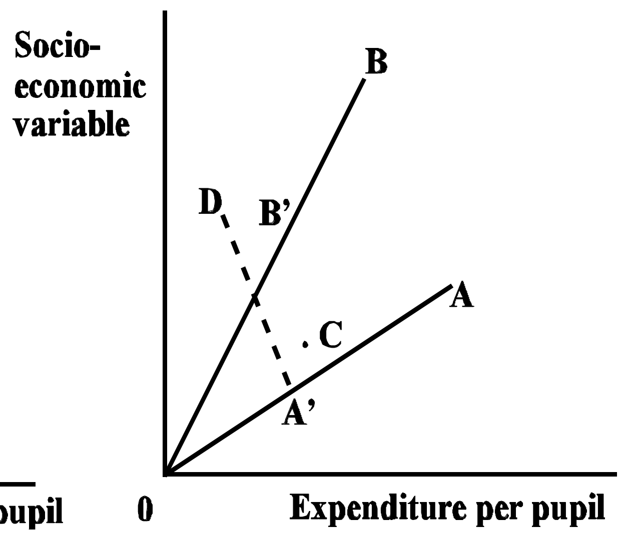

Figure 4b

In the above example, doubling the controllable variable of expenditure per pupil from D to B, or from C to A, results in a doubling of examination performance, for some changes in the uncontrollable variable. However, in the absence of homotheticity between all the inputs, both controllable and uncontrollable, full constant returns to scale with respect to proportionate changes in all the inputs do not prevail. Since the property of homotheticity depends upon which transformations of the variables are chosen, it is possible that selecting a different transformation of the socio-economic variables may succeed in achieving homotheticity. However, the use of DEA involves the problem that it does not generate an independent check on which of the schools A, B, C and D are actually efficient that is independent of the assumed homotheticity of the production structure. More generally, DEA can pose considerable difficulties in testing whether the underlying assumptions which DEA makes are fully appropriate, or alternatively are misspecified, for the particular application being studied. Some guide to the magnitude of the underlying problem will nevertheless be provided by the extent of the variation which occurs in the different DEA efficiency scores for a given school under different choices of transformation of the socio-economic variables, or under other choices of the model specification.

A further important implicit assumption of the standard formulation of DEA is that the production set, S, that defines the set of all technologically feasible input-output vectors (x,y), is *convex*. This assumption ensures that convex combinations of the input-output vectors of all DMUs in the comparison set are themselves feasible and can form the basis for computing the technical efficiency of the DMU whose unit is being assessed. For the case of multiple outputs, such as typically occurs in the education sector, convexity of S in particular requires that the rate of product transformation, given by minus the slope of the production possibility frontier between any two outputs, does not decrease as we move along the frontier.

A case where this condition may not hold arises if adequate care is not taken in the selection of the outputs of each educational institution. Thus one output of a school for its pupils at age 16 may be the number of its (points weighted) GCSE passes. Another may be the staying-on rate for its pupils, i.e. the number of its pupils who stay on for Sixth Form study. The production function for the first output, y_1, that of its GCSE passes, may be of the form:

$$y_1 = g (r_1 , á) \qquad (3)$$

being a function of the resources, r_1, that are devoted to this output, together with the vector á of its pupil characteristics. The staying-on rate, y_2 , may depend not only on the resources, r_2 ,that are directly devoted to encouraging an increase in this second output and on the characteristics of the pupil intake, á , but also upon the level of the schools' (points-weighted) GCSE passes, so that the production function for the second output is of the form:

$$y_2 = h (r_2 , y_1, á) \tag{4}$$

The rate of product transformation (RPT) between the two outputs, holding constant the total school resource use $r = r_1 + r_2$, is then given by:

$$RPT_{1,2} = - (dy_2 / dy_1)_{r = const} = (h_1 / g_1) - h_2 \quad \text{for } h_2 = \partial h / \partial y_1 \tag{5}$$

and where $g_1 > 0$ and $h_1 > 0$ are the partial derivatives of g and h respectively with respect to their resource input. Along the production possibility frontier, we then have:

$$(d (RPT_{1,2}) / dy_1)_{r = const} = - (d^2 y_2 / dy_1^2)_{r = const} \tag{6}$$

$$= - [h_{11} + (h_1 / g_1).g_{11} + g_1^2 . h_{22} - 2g_1.h_{12}] / g_1^2 \tag{7}$$

Under the assumption of diminishing marginal physical productivity of resources devoted to the two outputs, we have $g_{11} \equiv \partial^2 g / \partial r_1^2$ and $h_{11} \equiv \partial^2 h / \partial r_2^2$ both negative in sign. However, the signs of h_2 , $h_{12} \equiv \partial^2 h / \partial r_2 \partial y_1$ and $h_{22} \equiv \partial^2 h / \partial y_1^2$ are less clear-cut. A high level of GCSE passes may itself encourage more pupils to stay on into the Sixth Form in anticipation of being able to build further upon their academic success at GCSE. This itself will imply $h_2 > 0$ over this high range of GCSE passes. If this effect is more pronounced as GCSE passes rise further, we will have also $h_{22} > 0$. An increased level of success at GCSE may also boost the marginal productivity of resources devoted to encouraging a higher staying-on rate, implying $h_{12} > 0$ over this high range of GCSE passes.

However, a low level of GCSE passes may also encourage more pupils to stay on into the Sixth Form than would occur at moderate levels of GCSE passes, if those pupils with disappointing passes at GCSE (for their given pupil intake characteristics á) are seeking to retake GCSEs in the Sixth Form in order to attempt to remedy their previous lack of success. This itself will imply $h_2 < 0$ over this lower range of GCSE passes. Again, if this effect increases in absolute value as y_1 falls further, we will have $h_{22} > 0$ over this range. Falling GCSE passes may also boost the marginal productivity of resources devoted to encouraging a higher staying-on rate, implying $h_{12} < 0$ over this low range of GCSE passes.

Such values to h_{22} and h_{12} can clearly make the overall sign of (7) uncertain, given $g_{11} < 0$ and $h_{11} < 0$, and if large enough can cause (7) to become negative in sign over some range. The basic assumption of DEA, of a convex production set, will then be broken. The estimates of technical efficiency that DEA produces will them be unreliable, and possibly substantially so. Thus in Figure 5 below, the production possibility frontier is convex from above over the range of intermediate GCSE passes from C to F but becomes strictly concave from above between C and B over the lower range of GCSE scores. If the observed sample of schools is A, B and C, the technical efficiency coefficient which DEA produces for the school A based upon output maximisation will be OA/OE, where E is the point on the chord BC that lies on the ray OA through the origin. However, now that the assumption of convexity of the production set no longer holds, E is not in the feasible set, and A is actually on the efficiency frontier, and therefore should have a technical efficiency coefficient of 1.0 rather than OA/OE.

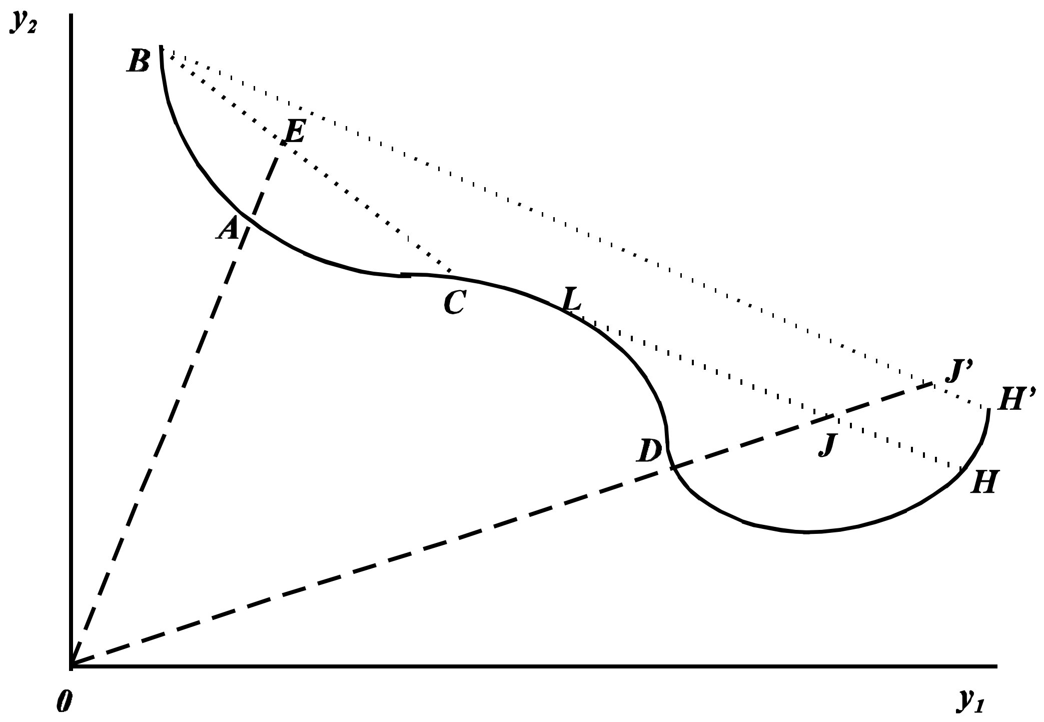

Figure 5

For schools in the high range of GCSE scores, the extent of the understatement of their true technical efficiency coefficient which DEA produces may be even greater. The rise in the value of h_2 as y_1 increases along the frontier may become large enough to cause RPT to become negative in (5). We then have an upward sloping section of the production possibility frontier in Figure 5 for high values of y_1 , where the direct effect of the increase in GCSE scores on boosting the staying-on rate in (4) is large enough now to overcome the indirect effect as resources are switched from y_2 to y_1. If the observed sample of schools is L, D and H, DEA would compute a technical efficiency coefficient of OD/OJ for school D, where J is on the chord between L and H and on the ray OD, even though D is perfectly efficient and on the efficiency frontier. If the observed sample of schools is extended to include B and H' in Figure 5, the extent of the understatement of D's true technical efficiency is even greater, with DEA producing an estimate of OD/ OJ' compared to a true value of 1.0.

A further difficulty which the application of DEA to education illustrates is that which arises if the effort, f , which each teacher exerts is an input variable which is not directly measured in the DEA study. In the case of a single output, y, of education (such as points-weighted GCSE scores), we may then have an underlying production function of the form:

$$y = G (N, K, f (á , N, K, y^T) , á) \qquad (8)$$

where N denotes the number of teachers employed in the school, K denotes capital inputs, á the characteristics of the pupil intake, and y^T is an output target that is set for the school. How the function f, which determines effort per teacher, behaves is then a critical element of the overall determination of the school's actual output y. For given values of á , N and K, we would expect f to increase with y^T over some initial range, so that over this range an increased target level of performance encourages greater effort per teacher. This may be particularly the case if a system of PRP is in place in which an actual output level that is equal to or greater than the target level is rewarded by significant additional performance payments to teachers, and this target level of output is achievable by reasonable levels of effort per teacher, given the values of N, K and á which the school faces.

However, as in Figure 6a, as the target level of performance, y^T, is increased further, it may well cause a decline in effort per teacher, as the target level is viewed to be less and less attainable and the financial pay-off from PRP less and less worthwhile for the additional effort per teacher that would be required to achieve an actual level of y at least as great as the target level, y^T. An optimal level of the output target, y^T, (if we ignore any additional financial costs which any associated system of PRP may itself impose) will be attained when f is at a maximum with respect to y', for given values of á , N and K in (8). When we substitute the optimal level $y^{T*} = y^{T*}(á , N, K)$ in (8), we obtain:

$$y = Q (N, K, á) = G (N, K, f (á , N, K, y^{T*}(á , N, K), á) \qquad (9)$$

For a given value of the target level of performance, y^T, the effort, f, per teacher in (8) will also vary with the number of teachers, N, in the school. With a high value to N, the given output target, y^T, may be attained with relatively little effort per teacher. As N is reduced from its initial high level, more effort per teacher is required, and may be forthcoming, as in Figure 6b, to attain the given target level of output, y^T. However, as N is further reduced, the given target level of output, and its contingent rewards from PRP, become less and less attainable, causing motivation to decline. Further reductions in the number of teachers, N, and consequent increases in class sizes, may cause further teacher demotivation, stress, absenteeism and increasing sickness rates due to stress.

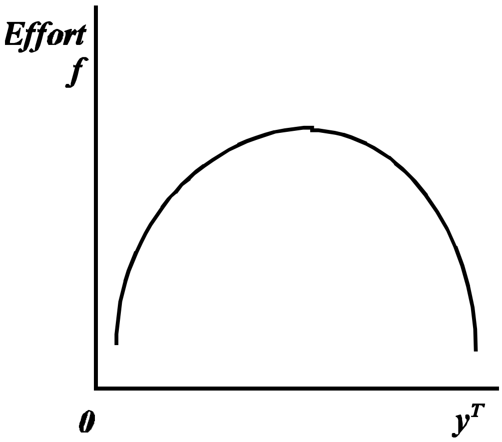

Figure 6a

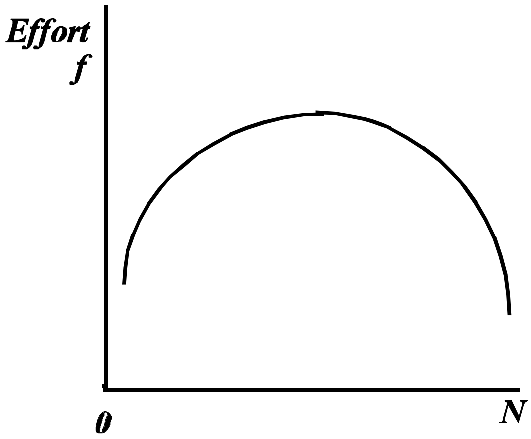

Figure 6b

Given that setting the output target optimally implies that $\partial f / \partial y^T = 0$, the second derivative of Q in (9) with respect to N is given by:

$$Q_{NN} = G_{NN} + G_{ff} \cdot f_N^2 + G_f \cdot f_{NN} + 2.\, G_{Nf} \cdot f_N - 2.G_f \cdot ((f_{NT})^2 / f_{TT}) \qquad (10)$$

where the subscript T denotes a derivative with respect to y^T .Convexity from the origin of the isoquants between N and K in the production of the educational output y in (9) can be shown to require that Q_{NN} is non-positive[36]. We may expect diminishing marginal physical productivity of N and f directly in q, and hence G_{NN} and G_{ff} both to be negative in (10). From Figure 6b we have here also $f_{NN} < 0$, with $G_f > 0$ in (9). However, we would expect that additional effort per teacher will raise the marginal physical productivity of additional numbers of teachers, implying $G_{Nf} > 0$. As in Figure 6a, f_N may be positive over an initial range. If raising the output target, y^T, has a diminishing impact on effort per teacher, then the last term of (10) will also be positive. Even if this last term is omitted because the output target y^T is held constant, rather than optimally varied as N is varied, (10) is overall of uncertain sign, for low values of N when f_N is positive. Convexity from the origin of the isoquants between the measured outputs N and K, holding á constant, is then not guaranteed, contrary to the assumptions of the standard model of DEA. The result of applying DEA as a performance measurement technique is this context may then again be an under-estimation by DEA of the actual coefficients of technical efficiency of schools in the sample that is under examination by reference to the inputs N, K and á .

Mis-estimating the value of the coefficient of technical efficiency may then lead to a number of problems in the context of performance management. The first is an over-estimate of the potential increases in educational outputs which the educational system as a whole can expect to achieve, from an examination of what the current performance of schools shows to be possible. The second is the setting of output targets for individual schools, as under the current *target-setting* process within LEA Education Development Plans[5], at an excessively high level. As in Figure 6a, this may adversely impact on teacher effort and the actual level of school performance achieved. The third problem is a potential disillusionment with performance reward systems, such as PRP, if some schools are set unattainable targets on the basis of under-estimates of their existing technical efficiency[6].

Knowledge, performance and resource management

The above discussion highlights the need for greater attention to be paid to the performance, reliability and potential pitfalls of performance measurement techniques themselves, and for sufficient care to be exercised in how they are deployed within performance management systems. The complex interactions between supply and demand considerations, and between labour and other inputs,and with characteristics of the pupil intake in producing multiple outputs, which are involved in the education system, mean that whilst a technique such as DEA has potentially much to offer in this context, a number of technical problems remain before full reliance can be placed upon its findings.

As with education itself, a key attribute of knowledge management in this context should be that of *learning* about the nature of the underlying problems and their practical significance. Rather than knowledge being a fixed entity which is then uncritically applied to real world systems, such as those of performance management, the process must be more *interactive*. Knowledge itself needs to be enhanced by the stimulation and questioning which the complexities of real world environments, such as education, can produce.

A further role for performance indicators in this context is then as a *trigger for further investigation*[1]. Rather than seeking to provide final answers to assessments of performance, performance measurement techniques in education and elsewhere can be deployed as part of a *quality control process* that utilises these techniques to trigger more detailed investigations into the complexities of an organisation's performance. Their use in such a process can explicitly recognise the risk, and relative costs, both of *Type 1 errors*, of investigating when the organisation is under sound existing management and its performance not capable of substantial improvement, and of *Type 2 errors*[37], of failing to investigate when the organisation and its management are capable of significant improvements. An existing process in primary and secondary education that has many potential attributes of a quality control system is that of OFSTED school inspection visits[2]. These involve substantial costs, not only from the direct cost of the inspectors' time, but also in terms of the *compliance costs* that are imposed upon the schools themselves in preparing for these visits. There is therefore a need to ensure that the expected benefits which these inspections achieve exceed the costs which they generate. The associated net benefit may be increased by deploying appropriate performance measurement techniques to better *target* inspection visits on schools where there is likely to be scope for *significant improvement* in performance from intervention.

The attainment of long-run improvements in performance is itself likely to involve a *continuing process of advice,* such as from local LEA advisors, rather than reliance simply upon infrequent 'big bang' OFSTED inspections. Such *qualitative* interaction [20], in conjunction with quantitative assessments from performance measurement techniques, may assist in enhancing our understanding of the strengths and limitations of these quantitative techniques. As a result of this feedback, these techniques may themselves be further improved. However, if knowledge itself becomes simply a *product for sale* by consultants in pre-packaged form in ways which cannot be readily questioned or challenged, there is a risk of false conclusions being drawn from over-confidence in its capabilities, and a lack of awareness of the limitations of existing performance measurement techniques.

A further important attribute of knowledge management should be the cost-effective deployment of *data* to enhance our understanding of the scope for improved performance, and for improved performance measurement techniques. As the author has argued in a recent report to the DfEE[19], there is considerable scope for a *national comparative database* to be established to the benefit of researchers, LEA advisers, school managers, OFSTED, the Audit Commission and central government. Such a database would include disaggregated data on school expenditure and resource use patterns, characteristics of the pupil intake, and value added outcome measures. Much of such data could be produced as low-cost by-products of well-designed routine management systems, such as computerised school time-tabling systems. Such data might shed considerable light on the pattern of deployment of teachers and other resources across different age groups and subject areas within the school, and on their resultant impact on school performance.

Knowledge of this kind is critical for improved resource management, both within the school, and in determining the details of optimal funding formulae for schools which can direct resources to where they are likely to be most educationally effective. In the current absence of such knowledge, there is no strong available research evidence to support the choice of key policy parameters in existing funding formulae. These include the differential age weights which are used in the computation of each school's age-weighted pupil numbers. These in turn make up at least 80 per cent of local school funding from its LEA and determine most of the relative funding balance between primary and secondary schools[19]. A suitable database would also assist in evaluating the effectiveness of policy innovations, such as the introduction of additional 20,000 classroom assistants and the increased use of educational technology in schools. Without carefully monitoring the effectiveness of these innovations, important knowledge will be unnecessarily lost, to the potential detriment of all participants in the education system.

The scope for making everyone better off, through Pareto improvements, is at the heart of the theoretical basis of DEA[35]. However, many existing applied DEA studies in education[3, 20, 38] and elsewhere have concentrated on computing each organisation's coefficient of *technical efficiency*, which focusses on the scope for performance improvements or cost reductions, whilst holding constant the existing proportionate input mix of the organisation. In the context of education, it can be argued that performance improvements are less likely to come from maintaining constant the existing proportionate mix between inputs. Instead, educational improvements may come through significantly changing the input mix, with much greater support for teachers from classroom assistants and new technology, in the form of IT, video systems and other capital investment to support learning. Such changes may become increasingly desirable as the relative price of able graduates tends to rise over time, and that of computer and other capital equipment tends to fall. New technology in service industries, as well as in manufacturing, is also likely to result in a growing pool of semi-skilled workers facing redundancy or forced early retirement, from whom many gifted classroom assistants might be recruited.

Given substantial recruitment problems[39] for teachers, particularly in areas such as mathematics and modern languages, improving teachers' marginal productivity and reducing the stress of teachers through greater supporting resources may not only directly improve school performance. They may also make possible improved pay and working conditions that enhance teacher recruitment, particularly of more able teachers, and thereby have further indirect positive influences on educational performance. Issues of price and allocative efficiency, which Farrell tended to downplay in his original seminal article[40] on the measurement of productive efficiency, become then of central importance to performance management in education. The need for care in the application of DEA is again underlined by the potential for the basic assumption of convexity (here of the set of capital and labour inputs capable of producing any given level of educational output) to be undermined by a strong degree of *complementarity* between capital and labour inputs[36], in which the marginal productivity of teachers is significantly raised by additional supporting resources, and vice versa.

Conclusion

Education provides an important area for the application of both performance measurement techniques and systems of performance management. It is also an area where great care must be taken at the interface between the two if dysfunctional outcomes are not to result. Performance measurement techniques, such as DEA and multivariate regression analysis, need themselves to be subject to a continuing *performance audit* to assess how far the potential complexities of the real world, which they may encounter in areas such as education, are adequately allowed for in their own assessment of performance. Such complexities, including those related to the interaction between educational performance, pupil and teacher motivation, stress, and recruitment, need also to be adequately addressed by performance management systems and their constituent parts, such as performance related pay. In addition, education illustrates the importance of an adequate linkage between performance management systems and issues of strategic management. Rather than concentrating mainly on issues of technical efficiency, performance management systems in education need also to address important strategic issues of price and allocative efficiency regarding the relative importance of different outputs of the educational system, and the scope for productively changing the input mix to provide greater support for individual teachers in the process of securing enhanced educational performance.

References

1, Mayston DJ (1985). Non-profit performance indicators in the public sector. *Financial Accountability and Management* **1**: 51- 74.

2, Office for Standards in Education (1995). *The OFSTED Handbook - Guidance on the Inspection of Secondary Schools*. Stationery Office: London.

3, Jesson D, Mayston DJ and Smith PC (1987). Performance assessment in the education sector: educational and economic perspectives. *Oxford Review of Education* **13**: 249 - 266.

4, Charnes A, Cooper W, Lewin A and Seiford L eds. (1994). *Data Envelopment Analysis - Theory, Methodology and Applications*. Klewer: Boston.

5, Department for Education and Employment (1999). *Departmental Report*. Cm 4202. Stationery Office: London.

6, Mayston DJ (1992). *School Performance Indicators and Performance Related Pay*. Assistant Masters and Mistresses Association: London.

7, H M Treasury (1998b). *Public Services for the Future: Modernisation, Reform, Accountability*, Cm 4181. Stationery Office: London.

8, HM Treasury (2000a). *Spending Review 2000: Public Service Agreements 2001 - 2004* . Cm 4808. Stationery Office: London.

9, Gillborn D and Youdell D (1999). *Rationing Education: Policy, Practice, Reform and Equity*. Open University Press: Buckingham.

10, Coopers and Lybrand Associates (1994). *Preventative Strategy for Young People in Trouble: Report*. Coopers and Lybrand: London.

11, Rawls, J.(1971). *A Theory of Justice*. Oxford University Press: London.

12, Atkinson, AB (1973). How progressive should income-tax be? In: Phelps, E (ed.) *Economic Justice*. Penguin: Harmondsworth, 386 - 408.

13, Department for Education and Employment (1999). *Secondary School Performance Tables 1999*. Http://www.dfee.gov.uk/performance/old_htm/schools_99.htm .

14, Assessment and Qualifications Alliance (2000). *GCE Interboard Statistics - Advanced Level 1990 - 1999.* AQA: Guildford.

15, Porter ME (1990). *The Competitive Advantage of Nations*. Macmillan: London.

16, Kay, J. (1993). *Foundations of Corporate Success*. Oxford University Press: Oxford.

17, Rierhoff H and Prais SJ (1997). *From School to Productive Work - Britain and Switzerland Compared*. Cambridge University Press: Cambridge.

18, Pedraja-Chaparro F, Salinas-Jimenez J and Smith P (1997). On the role of weight restrictions in data envelopment analysis. *J Productivity Analysis* **8**: 215 - 230.

19, Mayston DJ and Jesson D (1999). *Linking Educational Resourcing with Enhanced Educational Outcomes*. DfEE Research Report RR179. DfEE: London.

20, Mayston DJ and Jesson D (1988). Developing models of educational accountability. *Oxford Review of Education* **14**: 321 - 340.

21, Jesson D (1997). *Value Added Measures of School GCSE Performance*. Stationery Office: London.

22, Jesson D and Gray J (1991). Slants on slopes: using multi-level models to investigate differential school effectiveness and its impact on pupils' examination results. *School Effectiveness and School Improvement* **2**: 230 - 247.

23, HM Treasury (2000b). *Prudent for a Purpose: Spending Review 2000*. Cm 4807. Stationery Office: London.

24, Department of the Environment, Transport and the Regions (1998). *Modern Local Government: In Touch with the People*, Cm 4014. DETR: London.

25, Hanushek E (1986). The economics of schooling: production and efficiency in public schools. *J Econ Literature* **24**: 1141 - 1177.

26, Hanushek E (1995). Interpreting recent research on schooling in developing countries. *Wallis*

Institute of Political Economy Discussion Paper 3, University of Rochester.

27, Mayston DJ and Smith P (1990). Analysing the need to spend on education. *J Opl Res Soc* 41 (2): 125 - 131.

28, Mayston DJ (1996). Educational attainment and resource use: mystery or econometric misspecification? *Education Economics* 4 (2): 127 - 142.

29, Gujarati D (1995). *Basic Econometrics*, 3rd edn.. McGraw-Hill: New York.

30, Department of Education and Science (1983). *School Standards and Spending: A Statistical Analysis*. Statistical Bulletin 16/83. DES: London.

31, Department of Education and Science (1984). *School Standards and Spending: A Further Appreciation*. Statistical Bulletin 13/84. DES: London.

32, Mayston DJ (2000). The demand for education and the production of local public goods. University of York Discussion Paper.

33, Orme C and Smith P (1996). The potential for endogeneity bias in data envelopment analysis. *J Opl Res Soc* **47** (1): 73 - 83.

34, Shephard, RW (1970). *Theory of Cost and Production Functions*. Princeton University Press: New Jersey.

35, Norman M and Stoker B (1991). *Data Envelopment Analysis*. John Wiley: Chichester.

36, Henderson JM and Quandt RE (1980). *Microeconomic Theory*, 3rd edn. McGraw-Hill: New York.

37, Mittag H-J and Rinne H (1993). *Statistical Methods of Quality Assurance*. Chapman & Hall:

London.

38, Worthington A (1999). An empirical survey of frontier efficiency measurement techniques in education. *Education Economics*, forthcoming.

39, School Teachers' Review Body (1999). *Eighth Report*, Cm 4244, Stationery Office: London.

40, Farrell, MJ (1957). The measurement of productive efficiency. *J Roy Stat Soc* **120**: 11- 282.