Article:

Mayston, David John orcid.org/0000-0002-1157-3600 (2016) Convexity, quality and efficiency in education. Journal of the Operational Research Society. 446–455. ISSN: 1476-9360

https://doi.org/10.1057/jors.2015.91

Convexity, quality and efficiency in education

David J. Mayston

Department of Economics and Related Studies
University of York

Abstract: While Data Envelopment Analysis (DEA) has many attractions as a technique for analysing the efficiency of educational organisations, such as schools and universities, care must be taken in its use whenever its assumption of convexity of the prevailing technology and associated production possibility set may not hold. In particular, if the convexity assumption does not hold, DEA may overstate the scope for improvements in technical efficiency through proportional increases in all educational outputs and understate the importance of improvements in allocative efficiency from changing the educational output mix. The paper therefore examines conditions under which the convexity assumption is not guaranteed, particularly when the performance evaluation includes measures related to the assessed quality of the educational outputs. Under such conditions, there is a need to deploy other educational efficiency assessment tools, including an alternative non-parametric output-orientated technique and a more explicit valuation function for educational outputs, in order to estimate the shape of the efficiency frontier and both technical and allocative efficiency.

Keywords: Data envelopment analysis, quality, education, efficiency analysis, allocative efficiency.

1. Introduction

One of the most widely used techniques for analysing the efficiency of non-profit organisations in education and elsewhere is that of Data Envelopment Analysis (DEA) (see Emrouznejad *et al*, 2008; Johnes, 2015; De Witte and Lopez-Torres, 2015). In this context, DEA has many advantages through its ability to incorporate multiple outputs and multiple inputs into its determination of the efficiency scores of educational decision-making units (EDMUs), such as schools and universities. As a non-parametric frontier estimation technique, DEA also has the advantage of not requiring the prior specification of a specific functional form for the educational production function between educational inputs and outputs that maps out the frontier of the associated feasible set. However, an important assumption on which the conclusions of the standard models of DEA rest is that the technology, and associated production possibility set, is convex. In this paper we argue that care must be taken when this assumption may not be valid for many potential applications of DEA to assess educational efficiency. In Section 2, we examine how non-fulfilment of the convexity condition may lead to misleading conclusions on both the technical and the allocative efficiency of EDMUs. In Sections 3 - 5, we examine why the use of educational data in particular may lead to non-convexity. In Section 6, we examine the implications for efficiency assessments of not requiring the convexity assumption. Section 7 contains our conclusions.

2. The importance of the convexity assumption

The role of the convexity assumption in DEA's efficiency assessment can be seen most clearly in the output-orientated form of DEA developed by Banker *et al* (1984), which may be expressed in terms of the linear program:

$$\max \theta_j \ s.t. \ \lambda X \leq X_j, \theta_j Q_j - \lambda Q \leq 0, \lambda e = 1, \lambda \equiv (\lambda_1, ..., \lambda_n) \geq 0, e \equiv (1, ..., 1)' \qquad (1)$$

where X_j and Q_j are the input and output vectors respectively of the EDMU j, with $X \equiv (X_1,...,X_n)'$ and $Q \equiv (Q_1,...,Q_n)'$ from our sample of n individual EDMUs. (1) involves seeking the maximum possible proportional expansion θ_j in the existing output vector $Q_j = (Q_{j1},...,Q_{j\varpi})$ of EDMU j's ϖ different outputs from its existing input vector X_j based upon a comparison with a hypothetical EDMU. This is assumed to have an input vector that is a convex combination λ of the input vectors of the actual EDMUs in the sample and an output vector that is the same convex combination λ of the output vectors of the actual EDMUs in the sample. However, unless the feasible production possibility set is itself convex, there is no guarantee that this input-output combination of such a hypothetical EDMU, on which DEA's estimate of θ_j and its associated efficiency assessments are based, will actually be feasible. Moreover, as Halme *et al* (2014) note, managers may find such a hypothetical comparison unconvincing.

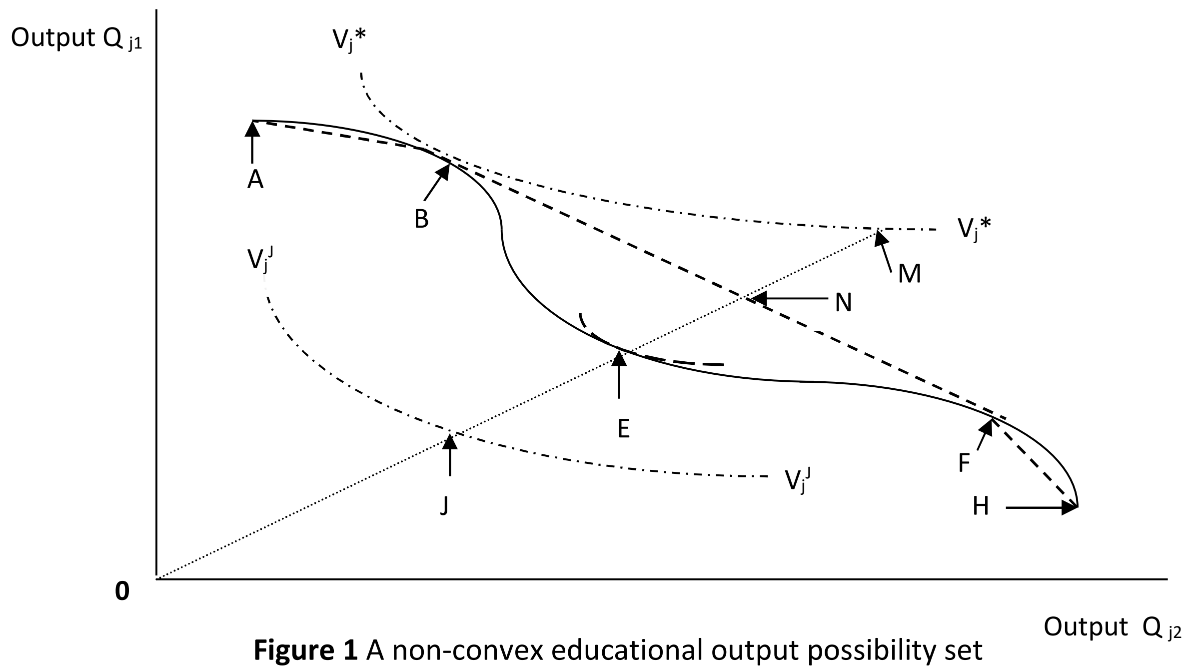

Figure 1 A non-convex educational output possibility set

Figure 1 illustrates a case in which the actual feasible set is non-convex, with a frontier given by the curve ABEFH between two educational outputs, with the level of inputs held constant in this simple example. The output vector that is achieved by the EDMU j in Figure 1 is given by the point J, which is strictly inside the feasible frontier ABEFH. The output-orientated form of DEA would compare the point J with a convex combination N of the feasible points B and F which are achieved here by other EDMUs in the observed sample, and where N lies on the same ray though the origin as point J.

The associated value of the coefficient of technical efficiency, ε_{TD}, of point J under DEA is OJ / ON. However, when the feasible set is non-convex, as in Figure 1, the point N may itself not be a feasible output vector. Instead, the true measure of technical efficiency, defined in terms of the proportional shortfall of its current output vector J compared to the maximum feasible proportional expansion in this output vector J at point E, would be here $\varepsilon_{TT} = OJ / OE$, which is strictly greater than DEA's coefficient of technical efficiency, ε_{TD}, OJ / ON. Even if it optimised a non-linear objective function with an indifference curve tangential at E in Figure 1, an EDMU that was actually on the efficient frontier at point E would erroneously be given by DEA a technical efficiency score of less than one, with an implied target for improvement at N that was not actually feasible, despite N having a potentially higher value than E under a linear objective function, such as a fixed-price revenue function.

While published DEA studies in education and elsewhere have concentrated predominantly on the assessment of *technical efficiency*, an important further direction in which DEA may yield biased efficiency assessments in the presence of non-convexity is in its assessment of *allocative efficiency*. Rather than focussing upon simply proportional improvements in the existing output vector, the concept of allocative efficiency seeks to assess which further improvements are feasible by changing the existing proportions in which the educational outputs are produced. One reason that output allocative efficiency has received much less attention in the educational literature than technical efficiency is that there are typically no simple market prices for the different educational outputs which an EDMU might produce. However, progress can be made in the assessment of the important issue of allocative efficiency if a *valuation function* of the form $V(Q_j)$ can be deployed to evaluate the relative value placed upon the different educational outputs. One useful property of such a valuation function in the context of allocative efficiency assessment is that of *homotheticity* (see Henderson and Quandt, 1980, p. 40), which implies that its iso-valuation curves are those generated by a valuation function which is homogeneous of degree one in the elements of the educational output vector Q_j. This in turn means that a relevant true measure of the allocative efficiency of the output vector J in Figure 1 is given by $\varepsilon_{AT} = (OE / OM)$, where M is the point on the ray through the origin on which J lies where it intersects the EDMU's actual *iso-valuation* curve, $V_i^* V_i^*$, that is assumed to be tangential to the feasible set at point B in Figure 1. The inverse of ε_{AT}, i.e. OM / OE, then provides a measure of the further improvements which can be made according to the homogeneous valuation function by changing the *educational output mix*, beyond those which can be achieved by improving its technical efficiency along the ray OE holding its existing output mix constant.

That DEA may overstate the existing allocative efficiency of the educational outputs of an EDMU in the presence of non-convexity is also illustrated in Figure 1, with DEA's measure of allocative efficiency given here by the ratio $\varepsilon_{AD} = (ON / OM)$. We then have:

$$\varepsilon_{AT} = (OE / OM) < (ON / OM) = \varepsilon_{AD} \ with \ \theta_{AT} \equiv \varepsilon_{AT}^{-1} > \varepsilon_{AD}^{-1} \equiv \theta_{AD} \tag{2}$$

with DEA's assessment, θ_{AD}, understating the true scope, θ_{AT}, for improvements in the value of educational outputs from changing the educational output mix from that along the ray OE to that at point B in Figure 1. That questions of output allocative efficiency and improving the educational output mix become more important in the presence of non-convexities than DEA recognises is consistent with the heightened need for educational institutions to make efficient choices regarding the mix of their educational outputs when there are sources of non-convexity in the production of such outputs.

3. Sources of non-convexity

In his seminar paper that led on to the development of DEA, Farrell (1957) acknowledged that "the whole method is based on the assumption of convexity". Yet, as Farrell (1959) himself stressed, "in the real world the relevant functions are often not convex" due to features such as indivisibilities and economies of scale. Scarf (1981) also stressed the need to acknowledge the existence of increasing returns to scale "which are implied by indivisibilites and other forms of nonconvexities in production". Moreover, Eaton and Lipsey (1997) have argued that "the mere existence of capital goods implied a fundamental nonconvexity in cost as a function of the stock of services embodied in an indivisible capital good, and therefore in the underlying production possibility set" and that "the non-convexities that arise from the once-and-for-all non-rivalrous nature of knowledge are pervasive and important". The non-rival nature of knowledge may indeed interact with the existence of capital goods in the case of higher education, if larger libraries and larger lecture theatres are able to benefit from economies of scale due to indivisibilities in a given range of library services and lecturer inputs. As well as knowledge once created and understood potentially exhibiting non-convexities in its wider use, the processes of creating knowledge through research, and of understanding it sufficiently to teach it well, may also involve non-convexities in the input-output space due to gains from specialisation in the development of the required human capital inputs. Productive research typically requires a high level of specialisation in reaching the frontiers of current knowledge and in developing the skills to make new contributions. Competition between research teams itself tends to push out the frontiers of knowledge and make more difficult the production of new contributions without greater inputs and more specialisation in increasingly technical directions. High quality teaching requires keeping up with specialist subjects, and their relation to other relevant issues, that are themselves changing over time.

The existence of such gains from specialisation implies non-convexity of the relationship between inputs and the efficient feasible output in any given quality direction. However, this in turn may have wider implications for non-convexity of the output possibility set defined by $P(x) = \{y : (x, y) \in T\}$ where T denotes the technology (i.e. the set of all technologically feasible input, output vectors) under which the EDMU operates. In our current context, the output vector $y = (m, q)$ includes the vector m of relevant quantities, such as student numbers at different stages of the educational process in a given institution, and the vector q of the EDMU's recorded quality achievements. For the sake of simplicity we will assume initially a single real resource input, whose total usage is held constant at some level x_0 in defining our output possibility set but whose allocation across different educational activities within the EDMU can be varied to change the educational output mix. We can then investigate the feasible output possibility set associated with a given total input x_0 given by:

$$P(x_0) = \{(q, m) : f(q, m, x_0) = x_0 - r(q, m) \geq 0\} \tag{3}$$

where $f(q, m, x_0) = 0$ defines an implicit multiple-output educational production function that maps out the production possibility frontier (PPF) for any given value of x_0, and which we assume can be decomposed into x_0 minus a function $r(q, m)$ that defines how the total resource input requirement varies with the quality vector q and the quantity vector m. Since the performance of educational institutions, such as schools and universities, is increasingly judged on the basis of the quality of their educational output, of particular interest is *the shape of the quality frontier*, holding m and x_0 constant. Even under the traditional assumption of microeconomics (see e.g. Henderson and Quandt, 1980) of differentiability of the production function, in which the slope of the PPF at any given point on the quality frontier corresponds to its *marginal rate of product transformation* $\Gamma_{kh} = (-dq_k / dq_h)$ between any two relevant quality scores holding constant m, x_0 and any other elements of q, a necessary condition (see Arrow and Enthoven, 1961) for convexity of $P(x_0)$ is that Γ_{kh} is non-decreasing as q_h is increased along the PPF, and hence that:

$$\chi_{kh} \equiv (d\Gamma_{kh} / dq_h) = (r_{hh} - 2r_{kh}(r_h / r_k) + r_{kk}(r_h^2 / r_k^2)) / r_k \geq 0 \; where \; r_k \equiv \partial r / \partial q_k, r_{kh} \equiv \partial r_k / \partial q_h \; etc \quad (4)$$

and where we assume that $r_k > 0$ and $r_h > 0$. Condition (4) in turn requires that:

$$(-r_{kh}) \geq 0.5[(-r_{kk}(r_h / r_k) - r_{hh}(r_k / r_j))] \quad (5)$$

When we include the quality of educational output within our efficiency analysis, we might indeed expect there to be increasing quality gains from specialisation and a greater focus of resources in particular directions, which would imply here that $r_{kk} < 0$ and $r_{hh} < 0$. The necessary condition (5) for convexity of the feasible set $P(x_0)$ then requires that any such gains from specialisation are offset by sufficiently large gains from the economies of scope associated with the cost complementarities (see Baumol *et al*, 1982, pp. 74-5) that negative r_{kh} terms in (5) reflect. Thus, it may be the case in universities that high quality research does indeed help to inspire high quality teaching, and that additional time spent teaching and preparing for teaching does generate some ideas for improved research activity, so that such production complementarities may well exist. However, the convexity condition (5) requires not simply that they exist but rather that they are sufficiently strong to offset the opportunity costs of the lost gains from specialisation due to a greater spreading of resources more thinly between the two activities. Whether or not this is the case is essentially an empirical question, rather than one which should assumed to be necessarily true, in the way the convexity assumption of DEA requires. Empirical evidence for a lack of any positive relationship between research and teaching quality is indeed claimed by Ramsden and Moses (1992). Similarly Marsh and Hattie (2002) conclude that “in contrast to the academic myth that research productivity and teaching effectiveness are complementary constructs, results of the present investigation – coupled with the findings of the Hattie and Marsh (1996) meta-analysis– provide strong support for the typical finding that the teaching-research relation is close to zero”. De Witte *et al* (2013) conclude that “once teaching time exceeds 20%, further increases in teaching duties seem to harm the overall academic performance. On the other hand, we observe that specialization in teaching and research correlates with better academic performance".

A breach of condition (5) also undermines the relevance of *time divisibility* that is used by Shephard (1970, p. 15) and Hackman (2008, p. 39) to justify convexity of the input possibility sets $L(y)$, of all the input vectors x that can produce at least as much output as the non-negative vector y, since spreading the available time input more evenly between two outputs undermines the gains from specialisation in a way that is not made up for by substantial production complementarities, even in the absence of any switching costs to move from one activity to another. Moreover, the associated cost function $C(y,w) = \{min\ wx : x \in L(y)\}$, where w is the input price vector for a general η_0-dimensional input vector x, is shown by Jacobsen (1970, p. 770) to be convex in the (ϖ - dimensional) output vector y if and only if the available technology $T = \{(x,y) \in \mathbb{R}_+^{\eta_0} \mathrm{x} \mathbb{R}_+^{\varpi} \mid y \in P(x)\}$ is itself convex. Non-convexity of T then implies non-convexity of the cost function in the output vector. In addition, Briec *et al* (2004) have shown that "in general, convex cost functions are never higher than non-convex cost functions" except in the case of a single output and constant returns to scale, so that "imposing convex cost targets may be excessively demanding when convexity is doubtful".

One multiple-output parametric production function that yields a non-convex output possibility set, and which can include both output quantity and quality variables, is that of the Cobb-Douglas form:

$$\prod_{k=1}^{\eta_1} q_k^{\alpha_k} \prod_{\kappa=1}^{\eta_2} m_\kappa^{\beta_\kappa} = A \prod_{\ell=1}^{\eta_0} x_\ell^{\gamma_\ell} \quad \textit{and hence } \Gamma_{kh} = (-dq_k / dq_h) = (\alpha_h q_k / \alpha_k q_h) \qquad (6)$$

where $\eta_1 > 0, \eta_2 > 0$, $\alpha_h > 0$ and $\alpha_k > 0$, with Γ_{kh} decreasing as q_h increases and q_k decreases along the PPF. The scope for its use to estimate the effectiveness of educational providers using a generalised form of Stochastic Frontier Analysis and a CES-valuation function for their outputs to assess allocative efficiency is discussed in Mayston (2015). A parametric production or cost function which provides a test of whether convexity does prevail is provided by the constant elasticity of transformation (CET) function:

$$r^{(j)} = a_o + (\sum_{h=1}^{\varpi} a_h y_{jh}^{\gamma_h})^{1/\rho} \quad where\ a_h > 0 \qquad (7)$$

used by Hasenkamp (1976a,b) for the case $a_o = 0, \gamma_h = \rho$ for $h = 1,...,\varpi$, and for which he found that the convexity condition of $\rho > 1$ failed to hold in his study of US railroad data. It is notable that a cross-section study by Izadi *et al* (2002) of the cost function of 99 UK universities using the above CET function also yielded estimates of the γ_h parameters strictly between zero and one, implying a non-convex iso-cost output possibility set, as in Baumol *et al* (1982, p. 461). There is therefore a need to allow for the possibility of non-convexity both in parametric and non-parametric applied production analysis.

4. Assessing educational quality

An important feature of the available data on the quality of educational output for schools and universities is that they typically result from *assigning grades* within the quality assessment process. For secondary schools in England, GCSE results achieved at grades A*- C have been a primary measure of the quality of their output, with much emphasis placed on the percentage of pupils who achieve 5 or more grades A*- C, including in English and mathematics. For universities in the UK, their research quality has been measured in terms of their submitted research outputs that fall within each of the grades 4*, 3*, 2*, 1* and unclassified (HEFCE 2010, 2015). UK university teaching quality is assessed by the percentage of student responses in the annual National Student Survey (NSS) that have been awarded grade 5, 4, 3, 2 or 1 according to the strength of their agreement with complimentary statements regarding their university department's teaching and associated provision (HEFCE, 2014).

The nature of the frontier of the feasible set between the quality scores in different directions facing an EDMU can be illustrated by the case of university research and teaching quality. We will assume that the resources devoted to one direction, such as research, could have spillover effects on the underlying quality achieved in another direction, such as teaching. In particular, we will denote by x_1 the resource expenditure on research per member of staff of any given EDMU and by x_2 its resource expenditure on teaching per student. The *underlying* quality y_{1i} of the ith assessed research output and the underlying quality y_{2i} of the ith assessed teaching episode are assumed to be given by:

$$y_{1i} = \bar{y}_1 + \varepsilon_{1i} \text{ where } \bar{y}_1 = \alpha_{11}x_1 + \alpha_{12}x_2, y_{2i} = \bar{y}_2 + \varepsilon_{2i} \text{ where } \bar{y}_2 = \alpha_{21}x_1 + \alpha_{22}x_2 \quad (8)$$

and where ε_{1i} and ε_{2i} are terms that reflect additional latent variations between each submission in individual ability and inspiration in research and teaching which impact on the underlying quality of each individual submission around the mean levels, $\bar{y}_1$ and $\bar{y}_2$ of the underlying quality given by the resource expenditures in Equation (8). The *assessed* quality of each submission, however, is a result of a *grading process* in which the grade awarded to the ith submission in direction k is given by:

$$\tilde{g}_{ki} = g_k \text{ if } \vartheta_{g_k+1} > y_{ki} \geq \vartheta_{g_k} \text{ for } k = 1,2 \quad (9)$$

with the grade hurdle $\vartheta_{g_k^{\infty}+1}$ for the highest grade g_k^{oo} assumed to be $+\infty$ and $\vartheta_{g_k^o}$ for the lowest grade g_k^o assumed to be $-\infty$. If ε_{1i} and ε_{2i} have independently normal frequency distributions with zero means and variances s_1^2 and s_2^2 respectively across the multiple individual submissions to the grading process, the *mean value of the assessed quality score* in direction k is given by:

$$q_k = Z_k(\bar{y}_k) \equiv \sum_{g_k} w_k(g_k)\Xi(\bar{y}_k, g_k) \text{ where } \Xi(\bar{y}_k, g_k) = [\Phi((\bar{y}_k - \vartheta_{g_k})/s_k) - \Phi((\bar{y}_k - \vartheta_{g_k+1})/s_k)] \quad (10)$$

and where Φ is the standardised normal cumulative distribution, $w_k(g_k)$ is the relative weight attached to the quality grade g_k in direction k, and $\Xi(\overline{y}_k, g_k)$ is the proportion of submissions in direction k awarded the quality grade g_k for a given value of $\overline{y}_k$. Equation (10) in turn implies that for $\Delta w_k(g_k) \equiv w_k(g_k) - w_k(g_k - 1) \geq 0$ for all g_k and $\Delta w_k(g_k) > 0$ for some g_k:

$$q'_k(\overline{y}_k) \equiv \partial Z_k / \partial \overline{y}_k = \sum_{g_k} \Delta w_k(g_k)\phi(Y_k)/s_k > 0 \quad where \;\; Y_k \equiv (\overline{y}_k - \vartheta_{g_k})/s_k \tag{11}$$

and

$$q''_k \equiv \partial q'_k / \partial \overline{y}_k = \sum_{g_k} \Delta w_k(\mathrm{g}_k) G(Y_k)/s_k^2 \;\; where\; G(Y_k) \equiv -Y_k\phi(Y_k) \tag{12}$$

where ϕ is the standardised normal density function. $\partial Z_k / \partial \overline{y}_k > 0$ in (11) and (10) imply inverse functions D_k such that $\overline{y}_k = D_k(q_k) \equiv Z_k^{-1}(q_k)$. Using (8), the associated cost function is given by:

$$c(q,m) = m_1 x_1 + m_2 x_2 = x_0 = z_1(m) D_1(q_1) + z_2(m) D_2(q_2) \tag{13}$$

where m_1 is the number of staff and m_2 the number of students of the EDMU, and where

$$z_1(m) \equiv (m_1\alpha_{22} - m_2\alpha_{21})/z_0, z_2(m) \equiv (m_2\alpha_{11} - m_1\alpha_{12})/z_0, z_0 \equiv (\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21}) \tag{14}$$

with $z_1 > 0$ and $z_2 > 0$ under the condition that $\partial c / \partial y_k > 0$ for $k = 1,2$ in (13). From (4), (11) – (14):

$$\chi_{12} = -(\mathrm{z}_2/\mathrm{z}_1)[q''_2 + q''_1(\mathrm{z}_2 / z_1)](q'_1 / (q'_2)^2) \tag{15}$$

The sign of χ_{12}, and hence whether or not the convexity condition (4) is broken, depends in (15) upon the behaviour, for both $k=1$ and $k=2$, of q''_k, and hence of the function $G(Y_k)$ at each point in the grading process for which $\Delta w_k(g_k) > 0$ in Equation (12). Figure 2 shows the strongly non-linear behaviour of the function $G(Y_k)$, with a steadily increasing positive value to $G(Y_k)$ as $\overline{y}_k$ increases up to the point where it is one standard deviation s_k short of the grade hurdle ϑ_{g_k}, and the associated values of Y_k and $G(Y_k)$ are minus one and +0.241971 respectively. $G(Y_k)$ then steadily declines from a positive to a negative value as the gap between ϑ_{g_k} and $\bar{y}_k$ passes through zero, with $G(Y_k)$ reaching a minimum of minus 0.241971 when $\bar{y}_k$ exceeds the grade hurdle ϑ_{g_k} by one standard deviation s_k, and the associated value of Y_k is plus one, before $G(Y_k)$ steadily increases in Figure 2 to approach zero.

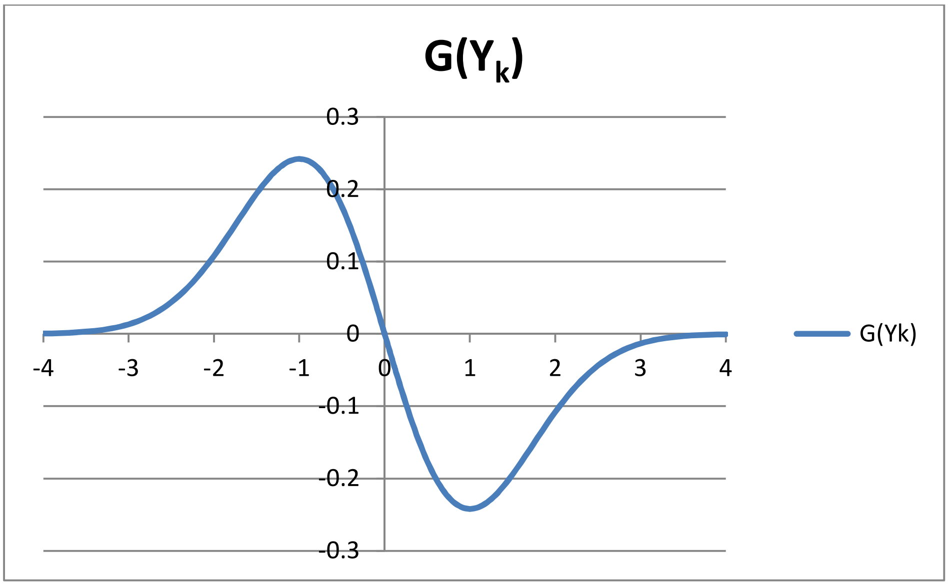

Figure 2 $G(Y_k)$ as a function of Y_k

As in Equation (12), the *rate of change of the marginal productivity* of increases in the mean value , $\bar{y}_k$, of the underlying quality from additional expenditure in Equation (8), in raising the *mean assessed quality score*, q_k, varies with $G(Y_k)$, and hence non-linearly with how close $\bar{y}_k$ is to the relevant quality hurdle. If the expected underlying quality $\bar{y}_k$ of the EDMU is just below the grade hurdle ϑ_{g_k} by close to one standard deviation s_k in both of the directions $k = 1,2$, and there is only one hurdle in each case for which $\Delta w_k(g_k) > 0$, we will have $q_1'' > 0$ and $q_2'' > 0$ in Equations (12) and (15), and hence the convexity condition $\chi_{12} \geq 0$ broken. Again this emphasises the importance of allocative efficiency, rather than simply technical efficiency, under such non-convexity. If there are insufficient slack resources to increase both $\bar{y}_1$ and $\bar{y}_2$ by more than one standard deviation s_k in each case, the EDMU would boost its assessed quality performance by choosing a more uneven policy of switching resources to boost $\bar{y}_1$ at the expense of $\bar{y}_2$, or vice versa, so that its expected underlying quality exceeds the relevant quality hurdle in at least one direction. Which one it should choose depends upon the relative payoffs from $\Delta w_1(g_1) > 0$ and $\Delta w_2(g_2) > 0$, with the extent of the improvement in its allocative efficiency dependent upon these relative payoffs and the non-proportional changes which it makes.

If there is only one grade hurdle in each direction and the EDMU reallocates its resources between x_1 and x_2 to further increase $\bar{y}_1$ and further reduce $\bar{y}_2$, while holding total cost constant, $G(Y_1)$, and hence q_1'', will at some point become negative in Figure 1 and in Equations (12) and (15); similarly $G(Y_2)$, and hence q_2'', will decline towards zero in its positive value. As a result, the necessary convexity condition $\chi_{12} \geq 0$ in Equation (15) may not be breached locally at all such points along the iso-cost frontier. This emphasises that while the associated output possibility set is here non-convex, as in Figure 1, its frontier is not everywhere concave from above. The existence of some concave sections to the frontier, as in Figure 1, will however make the use of DEA to assess the efficiency of individual EDMUs inappropriate when assessed quality variables are included in their outputs.

5. Multiple quality hurdles and the influence of league tables

The existence of possible multiple grade hurdles is illustrated by the case of the recent Research Excellence Framework (REF, 2014) exercise in UK universities and by its predecessor, the Research Assessment Exercise (RAE, 2008). The summation across these multiple grade hurdles that is involved in Equations (11) and (12) may well involve the expected underlying research quality $\bar{y}_1$ being above some lower quality hurdles, implying negative values to the associated $G(Y_1)$ terms in Figure 2, but below one or more higher quality hurdles, implying positive values to the associated $G(Y_1)$ terms in

Figure 2. Whether or not the overall weighted sum for q_1'' in Equation (12) is positive or negative will depend upon how far away $\bar{y}_1$ is from each such hurdle, and on the relative values of each respective $\Delta w_1(g_1)$ term. One notable feature of the relative weight placed upon achievements in successive assessed research quality grades from 1* through to 4* in the RAE and REF, in the determination of the associated QR research funding for each individual assessed EDMU by the Higher Education Funding Council for England (HEFCE), is that each successive increase $\Delta w_1(g_1)$ has itself been strictly increasing (see HEFCE, 2010, 2015), with $\Delta w_1(2^*) = 1$, $\Delta w_1(3^*) = 2$, and $\Delta w_1(4^*) = 6$ for the RAE and $\Delta w_1(2^*) = 0$, $\Delta w_1(3^*) = 1$, and $\Delta w_1(4^*) = 3$ for the REF. This itself implies a convex weighting function, with positive values to $G(Y_1)$ from the underlying expected research quality $\bar{y}_1$ being below a higher quality hurdle given more weight in (12) than a negative value to $G(Y_1)$ from $\bar{y}_1$ being at the same time above a lower quality hurdle, so that we may have $q_1'' > 0$ in Equations (12) and (15) over some range of values of $\bar{y}_1$.

When assessed teaching quality is included in the efficiency analysis, a notable feature of the relative weight that is placed upon successive grades from 1 to 5 in the NSS in published reports and performance indicators on the percentage of students who are "satisfied" (see e.g. HEFCE, 2014) is that $\Delta w_2(2) = \Delta w_1(3) = 0 = \Delta w_2(5)$ and $\Delta w_1(4) = 1$. The inclusion of only the NSS grades 4 and 5 as indicating that the student is "satisfied" means that when the assessed measure of teaching quality is the percentage of students who are "satisfied", there is no increase in this performance measure if a grade 3 rather than a grade 2 or a grade 1, or a grade 5 rather than a grade 4, is achieved. It is therefore the degree and sign of the difference between the expected underlying teaching quality y_2 and the quality hurdle associated with the boundary between grades 3 and 4 which determines the strength and sign of the relevant $G(Y_2)$ in Equation (12). If $\overline{y}_2$ falls short of this hurdle, we will have $q_2'' > 0$, with again the convexity condition $\chi_{12} \geq 0$ broken locally in (15) whenever $q_1'' > 0$ also holds.

Again issues of allocative efficiency become of considerable importance since the EDMU needs here to decide whether to boost their expected underlying teaching quality in order to increase the probability of grade 4 or 5 assessments, or to focus their available resources on boosting their expected underlying research quality to increase the probability of securing a higher research rating. Simply increasing both expected underlying qualities proportionately, to achieve increases in technical efficiency, in contrast may well prove to be a sub-optimal policy given the non-linearities which are involved in Equations (10) – (12), and the associated Figure 2. If their existing expected underlying teaching quality is between already well below the grade 4 hurdle, moderate improvements in it will unfortunately have little impact upon its overall expected assessed teaching quality score under the above weighting system. There is then more incentive for the EDMU to sacrifice even more teaching quality by concentrating its available resources more on improving its assessed research quality.

The powerful effect which the weighting system can have on an EDMU's management and policy choices in the presence of non-convexities is reflected also in the widespread use which has been made of the school performance indicator of the percentage of pupils who achieve 5 or more A* - C grades at the national GCSE examinations. This percentage similarly fails to give any additional positive credit for achieving higher grades within the A* - C range, and instead has given an incentive to schools to 'manage the margin' by focussing their resources upon pupils who are close to the grade C hurdle, rather than upon pupils who might excel towards grade A* performance. Fortunately some progress is being made through a recent proposal for a more refined point score system for 8 different grades at KS4 (DFE, 2014), though with pupil discreet grade improvements still playing a major part in the proposed Progress 8 performance measure from 2016 onwards.

An important additional pressure on the management of individual EDMUs can be their position in published national league tables of their assessed quality scores. The growth of *managerialism* within educational institutions (see e.g. Deem *et al*, 2007 and UNESCO, 2004) and of *greater competition* between EDMUs for able students based upon their published rankings (see e.g. DBIS, 2011 and Hazelkorn, 2011) significantly reinforces these pressures within individual EDMUs. The management objective of the EDMU may be described here as seeking to maximise a valuation function of the form:

$$V(R_1(q_1), R_2(q_2)) \ where \ R_k(q_\ell) = n\Psi_k(q_k) \ for \ k = 1,2 \tag{16}$$

where $R_k(q_k)$ is their rank from the bottom of the cumulative distribution $\Psi_k(q_k)$ of q_k across a given total number n of EDMUs in the comparison set. When we consider the possibility frontier facing the EDMU for changing its ranking in different quality directions, the slope of the relevant PPF between $R_1(q_1)$ and $R_2(q_2)$ holding total cost constant is given by:

$$\Gamma^o_{12} = -dR_1 / dR_2 = (dR_1 / dq_1)(-dq_1 / dq_2) / (dR_2 / dq_2) = \xi_1(q_1)\Gamma_{12} / \xi_2(q_2) \tag{17}$$

where $\xi_k(q_k)$ is the density function associated with $\Psi_k(q_k)$. We then have:

$$\chi^o_{12} \equiv d\Gamma^o_{12} / dq_2 = (\xi_1(q_1) / \xi_2(q_2))[-(\xi_{22}(q_2) / \xi_2(q_2))\Gamma_{12} - (\xi_{11}(q_1) / \xi_1(q_1))\Gamma^2_{12} + \chi_{12}] \tag{18}$$

If $\xi_k(q_k)$ is unimodal with a mode at q^o_k and a positive value to its slope $\xi_{kk} = \partial\xi_k / \partial q_k$ for $q_k < q^o_k$ and a negative value to its slope for $q_k > q^o_k$ for each $k = 1,2$, a given unit improvement in its assessed research quality score q_1 will give a potentially *much greater boost* to the EDMU's research ranking R_1 in (17) than otherwise, for any given reduction in its teaching ranking along the possibility frontier, if the EDMU is currently *close to the mode* of the associated population distribution for q_1 but distant from the mode of the population distribution for q_2. This in turn introduces another important source of non-linearity into the scope for increases in allocative efficiency, here with respect to the position of the EDMU in the relevant league tables. Moreover non-convexity of the associated possibility frontier may again arise here, with the convexity condition $\chi_{12}^{o} \geq 0$ not guaranteed to be fulfilled in (18) when $q_1 < q_1^o$ and/or $q_2 < q_2^o$, and hence $\xi_{11} > 0$ and/or $\xi_{22} > 0$, even if the local convexity condition $\chi_{12} \geq 0$ holds along the frontier between the q_k directly.

6. Non-parametric frontier analysis for non-convex production possibility sets

An output-orientated non-parametric technique that is consistent with the existence of both convex and non-convex regions of the efficient frontier can be generated by modifying the Free Disposal Hull (FDH) model proposed by Deprins *et al* (1984). This seeks the largest proportionate reduction in j's inputs that still involves at least as much of each input as an actual producer in the sample with which it is compared and which has achieved no less of each output as producer j, with the associated mixed-integer program (see Cooper *et al*, 2007):

$$min\ \tau_j\ \ s.t.\ \mu X \leq \tau_j X_j,\ \mu Q \geq Q_j,\ \mu u = 1,\ \mu_r \in \{0,1\}\ for\ each\ \ r = 1,...,n \tag{19}$$

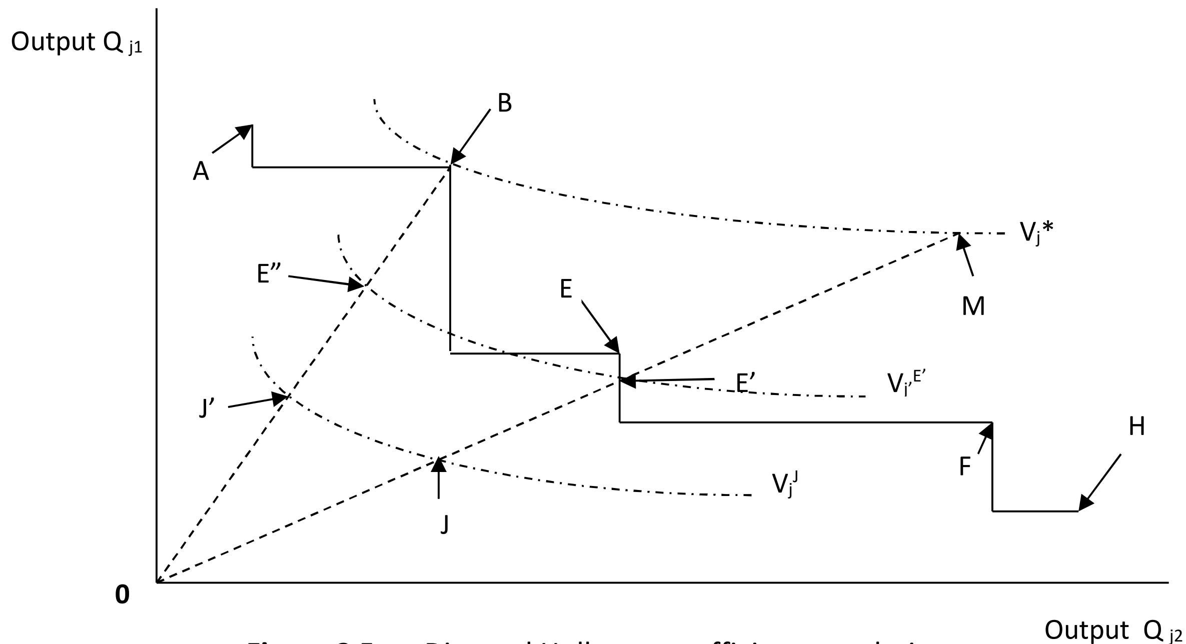

Figure 3 Free Disposal Hull output efficiency analysis

When inputs and outputs are both positive in value, we can transform (19) from an *input*- into an *output-orientated FDH model* by substituting the reciprocal of the output variables as inputs and the reciprocal of input variables as outputs in (19) and in the associated available software, to yield:

$$min\ \tau_j\ \ s.t.\ (1/Q_{\gamma k}) \leq \tau_j (1/Q_{jk}), (1/X_{\gamma \ell}) \geq (1/X_{j\ell})\ for\ each\ \ k=1,...,K\ and\ \ell=1,...,L \tag{20}$$

for an appropriate choice of the comparator producer γ. (20) in turn is equivalent to:

$$max\ \zeta_j\ s.t.\ \mu X \leq X_j, \zeta_j Q_j - \mu Q \leq 0, \mu u = 1, \mu_\gamma \in \{0,1\}\ for\ each\ \gamma = 1,...,n\ , with\ \ \tau_j = 1/\zeta_j \tag{21}$$

that generates an output-orientated alternative to the DEA program (1) which avoids DEA's requirement that the technology is convex. (21) results instead in a step-function form of the fitted efficient frontier through all the efficient points, such as A,B, E, F and H in Figures 1 and 3, that are actually achieved by individual producers, rather than the convex hull which DEA considers and which may include non-feasible convex combinations, such as point N in Figure 1. The value of τ_j in (21) provides a measure of technical efficiency of the EDMU j that equals the ratio OJ/OE' in Figure 3, and which has the property that its product with the corresponding measure of allocative efficiency OE'/OM equals the true overall measure of efficiency $\varepsilon_{OT} = OJ/OM$.

An alternative non-radial measure of technical efficiency suggested by Portela *et al* (2003) in another context would include here the additional slack $E'E$ that a comparison with the efficient point E indicates can be achieved in output Q_{j1}. However, once there are additional slacks in several directions, as in Tone (2001), defining a single such non-radial measure of technical efficiency raises the issue of the weight to be placed upon output increases in different directions. After examining alternative single radial and non-radial measures of technical efficiency under FDH, De Borger *et al* (1998) conclude that "unfortunately, none of these measures satisfies all of the desirable properties... with wide differences in the distributions of the efficiency scores and in the resulting correlations across alternative measures and orientations", despite the fact that "because the efficient subset is relatively small for the FDH reference technology, the choice among various efficiency measures is of crucial importance in measuring technical efficiency". However, rather than seeking a single measure of technical efficiency for FDH when significant slacks exist, an alternative approach adopted by Mayston (2014) is to compute both an overall measure based upon the maximum feasible proportional expansion in the existing output vector, and a set of individual measures that indicate the improvements which can be made in each individual output direction when the remaining slacks are included in addition to the overall maximum feasible radial expansion.

Use of an explicit valuation function would have the additional advantage of demonstrating what further movements along the efficiency frontier beyond point E, such as to point B in Figure 3, may be desirable. The reciprocal of the associated measure of allocative efficiency $\varepsilon_{Aj} = OE' / OM$ indicates the extent of the additional beneficial gains which can be made by changing the educational output mix from that at points J and E' to that at the optimal point B, with

$$\varepsilon_{Aj} = (OE' / OM) = V_j^{E'} / V_j^* = (OE'' / OB) \; and \; \varepsilon_{OT} = (OJ / OM) = V_j^J / V_j^* = (OJ' / OB) \qquad (22)$$

under a homothetic valuation function which is homogeneous of degree one in the elements of the educational output vector Q_j. Where an explicit valuation function is not available, Halme *et al* (2014) provide a method of incorporating into FDH binary preference information between pairs of existing outcomes to approximate the efficiency assessments that would be generated by a valuation function that is quasi-concave in outputs and quasi-convex.

While, as noted above, there are many DEA studies of efficiency in education, there are comparatively few that make use of FDH (exceptions are Afonso and St Aubyn, 2005; Oliveira and Santos, 2005; De Witte *et al*, 2010; Mayston, 2014), and even fewer studies that compare the results of deploying the two methods empirically. One which does is Mayston (2014), which found positive, though imperfect, Pearson and Spearman rank correlations between the efficiency score under the two methods, with DEA estimating only 34 per cent of the 50 UK university departments of economics to be fully efficient, in contrast to the 60 per cent which were found to be fully efficient under FDH. At the level of individual countries, Alfonso and St Aubyn (2005) found “very similar” DEA and FDH efficiency scores for their sample of 17 OECD countries using PISA average score results for 15 year-olds as outputs. In an earlier study of the wider range of public services provided by Belgian municipalities, Vanden Eeckaut *et al* (1993), however, found substantially higher efficiency estimates under FDH than DEA, concluding that “unless the convexity assumption can be given a strong a priori support – which is not the case with the data at hand – we see no reason to maintain it, and therefore reject the DEA results derived from it”. Similarly, in a study of Spanish municipalities, Balaguer-Coll *et al* (2007) found substantially higher technical efficiency estimates under FDH than DEA, with 69.8 per cent found to be fully efficient under FDH but only 7.7 per cent under DEA.

Elsewhere, Cummins and Zi (1998) concluded from their efficiency study of firms in the US life insurance industry that: “The distributional assumption imposed on the error term in the econometric models makes little difference with our data. Of much greater importance is the choice between econometric and mathematical programming methods, on the one hand, and whether to impose the convexity assumption in mathematical programming on the other” finding that DEA and FDH can yield “significantly different results”. Substantially higher values of technical efficiency were found using FDH rather than DEA by Wanke (2012) in his study of Brazilian airports, with many more container ports found to be technically efficient under FDH than DEA in the study by Cullinane *et al* (2005). The annual US federally managed fisheries capacity estimated by Walden and Tomberlin (2010) of 13.3 million pounds using DEA differs substantially from that of only 8.3 million pounds using FDH.

As Briec *et al* (2004) stress, in the case of large samples, “asymptotically, there is no reason for imposing convexity” since if the technology or the cost function “is truly convex, the FDH estimator converges to the true estimator" though with a lower convergence rate than for the convex estimator. More generally, Briec *et al* (2004) have suggested that the ratio between the input efficiency measures (cost function value) produced by DEA and FDH under similar returns to scale assumptions provides a non-parametric goodness of fitness test for the convexity of the underlying technology (cost function), which would imply that the wide divergences between the technical efficiency estimates produced under FDH and DEA call into question the empirical validity of the convexity assumption. More recently, Kneip *et al* (2015) have proposed an alternative more rigorous test for convexity based upon comparing the efficiency scores under FDH and DEA from different sub-samples of the dataset of DMUs, which they use to "test and soundly reject" DEA's assumption of convexity of the production set for the US banking industry.

7. Conclusions

Particularly when published measures of the assessed output quality of educational organisations, such as schools and universities, are included in the efficiency analysis, DEA's underlying assumption of convexity of the associated feasible set may not hold, so that care must be taken in the choice of non-parametric technique to estimate the associated scope for improvements in technical efficiency. At the same time, greater recognition may be needed of the scope for improvements in allocative efficiency from non-proportional changes in the educational output mix. An assessment of allocative efficiency itself requires a clarification of the valuation function which the institution places upon the volume and assessed quality of its educational outputs. When educational institutions are under managerial pressure to perform well according to their rankings in published league tables, the scope for non-convexities may increase, further boosting the importance of allocative efficiency and the need for more stark choices to be made by individual EDMUs between different assessed output quality variables. The shape of the output possibility frontier which maps out the trade-offs that an EDMU may face between the different assessed output quality variables at different points on the frontier may be better revealed by relaxing the convexity assumption, as in the above output-orientated FDH model. At the same time, national policy makers need to re-assess whether the grading and weighting systems which result in the management choices which individual EDMUs face along this frontier of assessed outcomes are consistent with wider educational goals.

References

Afonso A and St. Aubyn M (2005). Non-parametric approaches to education and health expenditure efficiency in the OECD. *Journal of Applied Economics* **8**(2): 227-246.

Arrow K and Enthoven A (1961). Quasi-concave programming. *Econometrica* **29**(4): 779 – 800.

Balaguer-Coll M, Prior D and Tortosa-Ausina E (2007). On the determinants of local government performance: a two-stage nonparametric approach. *European Economic Review* **51**(2): 425-451.

Banker R, Charnes A and Cooper W (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science* **30**(9): 1078-1092.

Baumol W, Panzar J and Willig R (1982). *Contestable Markets and the Theory of Industry Structure*. Harcourt Brace Jovanovich: New York, NY.

Briec W, Kerstens K and Vanden Eeckaut P (2004). Non-convex technologies and cost functions: definitions, duality and nonparametric tests of convexity. *Journal of Economics* **81**(2): 155-192.

Cooper W, Seiford L and Tone K (2007). *Data Envelopment Analysis*, 2nd edn. Springer: New York, NY.

Cullinane K, Song D-W and Wang T (2005). The application of mathematical programming approaches to estimating container port production efficiency. *Journal of Productivity Analysis* **24**(1): 73-92.

Cummins JD and Zi H (1998). Comparison of frontier efficiency methods: an application to the US life insurance industry. *Journal of Productivity Analysis* **10** (2): 131-152.

De Borger B, Ferrier G and Kerstens K (1998). The choice of a technical efficiency measure on the Free Disposal Hull reference technology: a comparison using US banking data. *European Journal of Operational Research* **105**(3): 427-446.

Deem R, Hillyard S and Reed M (2007). *Knowledge, Higher Education, and the New Managerialism*. Oxford University Press: Oxford.

Department for Education (2014). *Progress 8 School Performance Measure*. DFE: London.

Department of Business, Innovation and Skills (DBIS) (2011*). Higher Education: Students at the Heart of the System*. Cm 8122. HMSO: London.

Deprins D, Simar L and Tulkens H (1984). Measuring labor-efficiency in post offices. In: Marchand M, Pestieau P and Tulkens H (eds) *The Performance of Public Enterprises - Concepts and Measurement*, North-Holland: Amsterdam, pp. 243-267.

De Witte K, Thanassoulis E, Simpson G, Battistri G and Charlesworth-May A (2010). Assessing pupil and school performance by non-parametric and parametric techniques. *Journal of the Operational Research Society* **61**(8): 1224-1237.

De Witte K, Rogge N, Cherchy L and van Puyenbroek T (2013). Economies of scope in research and teaching: a non-parametric investigation. *Omega* **41**(2): 305-314.

De Witte K and Lopez-Tores L (2015). Efficiency in education: a review of literature and a way forward, mimeo.

Eaton C and Lipsey R (1997). *On the Foundations of Monopolistic Competition and Economic Geography.* Edward Elgar Press: Cheltenham, pp. x and 110.

Emrouznejad A, Parker B and Tavares G (2008). Evaluation of research in efficiency and productivity: a survey and analysis of the first 30 years of scholarly literature in DEA. *Socio-Economic Planning Sciences* **42**(3): 151-157.

Farrell M (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society* Series A **120**(3): 253-290.

Farrell M (1959). The convexity assumption in the theory of competitive markets. *Journal of Political Economy* **67**(4): 377-391.

Hackman S (2008). *Production Economics: Integrating the Microeconomic and Engineering Perspectives.* Springer: Berlin.

Halme M, Korhonen P and Eskelinen J (2014). Non-convex value efficiency analysis and its application to bank branch sales evaluation. *Omega* **48**: 10-18.

Hasenkamp G (1976a). A study of multiple-output production functions. *Journal of Econometrics* **4**(3): 253-262.

Hasenkamp G (1976b). *Specification and Estimation of Multiple-Output Production Functions*. Springer: Berlin.

Hattie J and Marsh H (1996). The relationship between research and teaching: a meta-analysis. *Review of Educational Research* **66**(4): 507-542.

Hazelkorn E (2011). *Rankings and the Reshaping of Higher Education*. Palgrave Macmillan: Basingstoke.

Henderson J and Quandt R (1980). *Microeconomics Theory – A Mathematical Approach*, 3rd edn. McGraw-Hill: New York, NY.

HEFCE (2010). *How HEFCE Allocates Its Funding*. HEFCE: Bristol.

HEFCE (2014). *UK review of the provision of information about higher education: National Student Survey results and trend analysis 2005-2013.* HEFCE: Bristol.

HEFCE (2015). *How We Fund Research*. http://www.hefce.ac.uk/rsrch/funding/mainstream/ accessed 8th July 2105.

Izadi H, Johnes G, Oskrochi R and Crouchley R (2002). Stochastic frontier estimation of a CES cost function: the case of higher education in Britain. *Economics of Education Review* **21**(1): 63-71.

Jacobsen S (1970). Production correspondences. *Econometrica* **38**(5): 754-771.

Johnes J (2015). Operational research in education. *European Journal of Operational Research* **243**(3): 683-1028.

Kneip A, Simar L and Wilson P (2015). Testing hypotheses in nonparametric models of production. *Journal of Business and Economic Statistics*, forthcoming, DOI:10.1080/07350015.2015.1049747.

Marsh H and Hattie J (2002). The relation between research productivity and teaching effectiveness: complementary, antagonistic, or independent constructs? *Journal of Higher Education* **73**(5): 603-641.

Mayston DJ (2014). Effectiveness analysis of quality achievements for university Departments of Economics. *Applied Economics* **46**(31): 3788-3797.

Mayston DJ (2015). Analysing the effectiveness of public service producers with endogenous resourcing. *Journal of Productivity Analysis* **44**(1): 115-126.

Oliveira MA and Santos C (2005). Assessing school efficiency in Portugal using FDH and bootstrapping. *Applied Economics* **37**(8): 957-968.

Portela MCS, Borges P and Thanassoulis E (2003). Finding closest targets in non-oriented DEA models: the case of convex and non-convex technologies. *Journal of Productivity Analysis* **19**(2-3): 251-269.

Ramsden P and Moses I (1992). Associations between research and teaching in Australian higher education. *Higher Education* **23**(3): 273-295.

Scarf H. (1981). Production sets with indivisibilities Part II: the case of two activities. *Econometrica* **49**(2), 395-423.

Shephard R (1970). *Theory of Cost and Production Functions.* Princeton University Press: Princeton.

Tone K (2001). A slacks-based measure of efficiency in data envelopment analysis. *European Journal of Operations Research* 130 (3): 498-509.

UNESCO (2004). *Managerialism and Evaluation in Higher Education.* UNESCO Forum Occasional Paper Series Paper No. 7. UNESCO: Paris.

Vanden Eeckaut P, Tulkens H and Jamar M-A (1993). Cost efficiency in Belgian municipalities. Chapter 12 in *The Measurement of Productive Efficiency*, Fried H, Lovell CAK and Schmidt S (eds). Oxford University Press: Oxford.

Walden J and Tomberlin D (2010). Estimating fishing vessel capacity: a comparison of nonparametric frontier approaches. *Marine Resource Economics* **25**(1): 23-36.

Wanke P (2012). Efficiency of Brazil's airports: evidences from bootstrapped DEA and FDH estimates. *Journal of Air Transport Management* **23**: 47-53.