THE UNIVERSITY of York

Discussion Papers in Economics

No. 11/25

Analysing the Research and Teaching Quality Achievement Frontier

By

DAVID MAYSTON

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

Analysing the Research and Teaching Quality Achievement Frontier

DAVID MAYSTON*

Abstract

The paper analyses the nature of the achievement possibility frontier between research and teaching quality in higher education under a system of quality evaluation by reference to quality grades. It finds several important reasons why the associated feasible set is likely to be non-convex, particularly in the presence of endogenous resource inputs, and hence where the assumptions of the widely used non-parametric frontier performance evaluation technique of Data Envelopment Analysis no longer hold. The paper therefore investigates the use of the alternative Free Disposal Hull technique, and compares the results of deploying these techniques to the performance and efficiency evaluation of UK Departments of Economics.

Keywords: Research and teaching quality, Higher education, Departments of Economics, Data Envelopment Analysis (DEA), Free Disposal Hull

JEL Classification: A29, C65, D24, H42, I21, I23

____

* Department of Economics and Related Studies, University of York; dm3@york.ac.uk

The provision of detailed statistical data by the UK Higher Education Statistical Agency is gratefully acknowledged.

1. Introduction

The assessed quality of both the research and the teaching carried out in university departments is of increasing importance to their perceived academic success and continuing vitality. This is particularly so in a higher education system, such as that in the UK, where there are now in operation formalised mechanisms for such quality assessment at the level of individual academic departments. The resultant evaluations of their research and teaching quality are likely to become of even greater significance with increased competition between universities for research funding, that is itself under pressure from tighter public expenditure constraints, and for students facing rising tuition fees seeking greater perceived value for money from investment in their own human capital. An increased emphasis on assessed teaching quality is illustrated by the recent UK Government White Paper (DBIS, 2011) *Higher Education: Students at the Heart of the System*, which envisages "a more dynamic sector in which popular institutions can grow and where all universities must offer a good student experience to remain competitive" (ibid, p. 5), with published measures of student satisfaction with different aspects of their higher education teaching forming a key part of the desired process of "well-informed students driving teaching excellence" (ibid, p. 28).

With individual academic Departments in the UK already under pressure to maintain or improve the research quality assessments that they received in the most recent Research Assessment Exercise (HEFCE, 2008) in the forthcoming Research Excellence Framework exercise to be completed in 2014 (HEFCE, 2011), important questions arise as to the joint feasibility of improvements in all the desired directions of teaching and research under tight resource constraints, and the nature of the frontier of the feasible set in this context. While techniques, such as Data Envelopment Analysis (DEA) (see e.g. Cooper, Seiford and Tone, 2007) and Stochastic Frontier Analysis (SFA) (see e.g. Kumbhakar and Lovell, 2000), exist for attempting to empirically assess the extent of such feasible improvements, there is a need to examine their appropriateness for analysing the extent of improvements that are feasible in the assessed quality of both research and teaching for individual academic Departments. In this paper, we pursue this examination firstly using a theoretical analysis and secondly by analysing empirical data for higher education Departments of Economics in the UK.

2. The Achievement Frontier Between Research and Teaching Quality

Let $\Xi \equiv \{1,...,s\}$ be the set of academic Departments in which we are interested. Each Department $j \in \Xi$ is assumed to be assessed on its quality of output from its activities in a number of directions. Direction $k = 0$ will designate its research activity that is assessed under a process similar to the RAE and which is assumed here to focus on the quality of the research output of its research-active members of its academic staff. Directions $k = 1,...,K$ will designate different aspects of its teaching activities that are assessed under a process similar to the NSS that focuses on responses per student to questionnaires on their degree of satisfaction with the different aspects of teaching within the Department. Each Department makes use of academic staff time and supporting resources in these different directions. The index of real resource r_{kj} devoted to activity k by Department j is assumed to be positively related to its use of academic staff time t_{kj} and supporting resources z_{kj} via a Cobb-Douglas production function of the form:

$$r_{kj} = A_k x_{kj}^{\alpha_k} z_{kj}^{1-\alpha_k} \quad where\ x_{kj} = a_{kj} t_{kj} \quad for\ k = 0,...,K \tag{2.1}$$

where a_{kj} is a parameter which reflects the ability level of the academic staff which Department j deploys in direction k, and $\alpha_k < 1$ and A_k are positive constants. In the case of research, r_{0j}, t_{0j} and z_{0j} are measured in terms of resources per research-active member of the Department, and given a value of zero if there are no research-active members of the Department. In the case of teaching, r_{kj}, t_{kj} and z_{kj} are measured in terms of resources per student. Department j can purchase supporting resources at a price of ω_2 per unit of z_{kj}, and can hire academic staff time at a wage ω_{1kj} per unit of x_{kj}, so that the total wage paid by Department j for activity k increases with the time spent on this activity and with the ability level of the staff involved. Its cost function for devoting the level of real resource r_{kj} to activity k is therefore given by:

$$C^*_{kj}(r_{kj}, \omega_{1kj}, \omega_2) \equiv [\min_{x_{jk}, z_{kj}} C_{kj} = \omega_{1kj} x_{kj} + \omega_2 z_{kj} : r_{kj} - A_k x_{kj}^{\alpha_k} z_{kj}^{1-\alpha_k} = 0\,] = r_{kj} (\omega_{1kj} / \alpha_k)^{\alpha_k} (\omega_2 / (1-\alpha_k))^{1-\alpha_k} / A_k \qquad (2.2)$$

Synergies may exist between its different activities, so that a high level of resources devoted to research by Department j may succeed in boosting the extent to which students find its teaching intellectually stimulating. Similarly, more resources devoted to making the Department's teaching intellectually stimulating, by keeping its courses up to date in their content, may stimulate increased research output by the Department. Negative synergies are also, however, possible, with a greater research orientation of the Department possibly resulting in tougher standards for awarding degree classifications within the Department, that may result in lower levels of student satisfaction in some directions. The extent of these synergies may be represented by the elements in the matrix B for determining the overall value v_{kj} of the real resources that impact upon the quality assessment of activity k within Department j, with:

$$v_{0j} = \sum_{h=0}^{K} b_{0h} r_{hj} \;\; and \;\; v_{kj} = \sum_{h=0}^{K} b_{kh} r_{hj} \;\; for\; k = 1, ..., K, \; where\; B \equiv [b_{kh}\,], b_{kk} > 0 \;\; for\; k, h = 0, ..., K \qquad (2.3)$$

with b_{kh} for $k \neq h$ positive, zero or negative according to the direction of the synergy between activities k and h
.

The quality of any given item of research that is produced by a research-active member of academic staff in Department j is assumed to be given by:

$$y_{kj} = v_{kj} + \varepsilon_{kj} \qquad (2.4)$$

for $k = 0$, which contains a stochastic element ε_{0j} that is assumed to be iid $N(0, \sigma^2_{01})$, reflecting the serendipitous nature of research outcomes. Similarly, the underlying quality of teaching within Department j is assumed to be given by (2.4) for $k = 1, ..., K$, where the stochastic elements ε_{kj} are assumed to be iid $N(0, \sigma^2_{k1})$, and to reflect influences upon the underlying teaching quality that are unrelated to resourcing levels.

The assessment of the quality of research that is produced is assumed to be via an evaluation process of submitted publications and other research outputs according to their assessed quality in a Research Assessment Exercise (RAE), such as was completed for UK universities in 2008. The assessment of the quality of teaching by Department j in direction k is assumed to be also via questionnaires to students that rank the quality of teaching in direction k into a number of discrete grades.

Specifically, we will assume that the underlying quality y_{kj} of Department j's activity in direction $k = 0,...,K$ will be given an assessed quality of $y_{kj}^* = g_k$ (where a higher value of g_k indicates a higher grade) according to whether it falls within the grading thresholds for grade g_k given by:

$$y_{kj}^* = g_k \quad if\ \gamma_{k,g_k} + \varepsilon'_{kj} \le y_{kj} < \gamma_{k,g_k+1} + \varepsilon'_{kj} \quad where\ \gamma_{k,g_k+1} > \gamma_{k,g_k} \tag{2.5}$$

γ_{g_k} is here the expected level of the quality in direction k that must be achieved to secure at least a grade g_k rating. However, in addition to this expected level in (2.5), there exists a degree of uncertainty, as reflected in the stochastic term ε'_{kj}, surrounding the precise quality thresholds which are applied within the assessment processes for mapping any given underlying quality in direction k, as given by the continuous variable y_{kj}, into a discrete assessed quality grade g_k. For the sake of simplicity, we assume here that this uncertainty affects all threshold margins for direction k equally, with the ε'_{kj} iid $N(0,\sigma_{k2}^2)$. (2.4) - (2.5) imply for $k = 0,...,K$:

$$y_{kj}^* = g_k \ \ if\ \ \gamma_{k,g} - v_{kj} \le \varepsilon''_{kj} < \gamma_{k,g+1} - v_{kj}\ where\ \varepsilon''_{kj} \equiv \varepsilon_{kj} - \varepsilon'_{kj}\ \ is\ \ N(0,\sigma_k^2)\ and\ \sigma_k^2 = \sigma_{k1}^2 + \sigma_{k2}^2 \tag{2.6}$$

The different grades of assessed quality in direction k are in turn assumed to be weighted by the assessment process, according to a weighting function $w_k(g_k)$, to produce an overall measure of the quality of activity k. In the case of the RAE, the weights attached by HEFCE (2010a) within its associated QR funding formula are 9 for a grade of 4*, 3 for a grade of 3*, 1 for a grade of 2* and 0 for 1* and unclassified, so that $w_0(4^*) = 9, w_0(3^*) = 3, w_0(2^*) = 1, w_0(1^*) = 0$. In the case of the NSS, grades 1 to 5 correspond to the assessed strength of student agreement with statements, such as "staff have made the subject interesting", on different aspect of teaching, with grade 1 corresponding to "definitely disagree", grade 2 to "mostly disagree", grade 3 to "neither agree nor disagree" , grade 4 to "mostly agree" and grade 5 to "definitely agree". The weighting function used in several published reports on the results of the NSS (see e.g. HEFCE, 2010b) is that of 1 for grades 4 and 5, with zero otherwise. We will assume more generally that:

$$w_k(g_k) \ge w_k(g_k - 1)\ for\ all\ g_k = 1,...,\hat{g}_k \ \ and\ w_k(g_k) > w_k(g_k - 1)\ for\ some\ g_k = 1,...,\hat{g}_k \tag{2.7}$$

where $\hat{g}_k$ is the highest classified grade and 0 is the lowest classified grade for direction $k = 0,...,K.$

The resultant expected assessed quality of output of Department j in direction k is:

$$q_{kj} = \sum_{g_k=0}^{\hat{g}_k} w_k(g_k)[N((\gamma_{k,g_k+1} - v_{kj})/\sigma_k) - N((\gamma_{k,g} - v_{kj})/\sigma_k)] \tag{2.8}$$

for each assessed item, where N is the cumulative normal distribution function. We will assume that the total number of publications which Department j submits for assessment, and the number of students in Department j, are both large enough to make the actual average assessed quality of output in each direction arbitrarily close to its expected value. For any given set of values of the parameters $\gamma_{g_k}, g_k, \sigma_k$ etc, (2.8) implies that for $k = 0,...,K$ the q_{kj} are functions of the form:

$$q_{kj} = f_k(v_{kj}) \;\; with \;\; f_k'(v_{kj}) \equiv \partial f_k(v_{kj}) / \partial v_{kj} = -\sum_{g=0}^{\hat{g}_k} w_k(g)[\phi((\gamma_{k,g+1} - v_{kj})/\sigma_k) - \phi((\gamma_{k,g} - v_{kj})/\sigma_k)]/\sigma_k \quad (2.9)$$

where ϕ is the normal density function. If we assume that the highest quality threshold $\gamma_{k,\hat{g}_k+1}$ in (2.7) – (2.9) goes to $+\infty$ and that the lowest quality threshold $\gamma_{k,0}$ goes to $-\infty$, so that evaluations are possible over the entire spectrum of underlying performance, we have from (2.5), (2.8) – (2.9):

$$f_k'(v_{kj}) = \sum_{g_k=1}^{\hat{g}_k} \Delta w_k(g_k)\phi((\gamma_{k,g_k} - v_{kj})/\sigma_k)/\sigma_k > 0 \;\; for\ all\ v_{kj} \;\; where\ \Delta w_k(g_k) \equiv w_k(g_k) - w_k(g_k - 1) \quad (2.10)$$

From (2.3), (2.9) and (2.10) there exists an inverse function $\upsilon_k = f_k^{-1}(q_{kj})$ such that the level of real resource inputs in each direction that is required to achieve an expected quality of output of q_{kj} in each direction is given by :

$$r_{hj} = \tau_{h0}\upsilon_0(q_{oj}) + \sum_{k=1}^{K} \tau_{hk}\upsilon_k(q_{kj}) \equiv \chi_{hj} \quad for\ each \;\; h = 0,...,K \;\; where\ [\tau_{hk}] \equiv \Gamma \equiv B^{-1} \quad (2.11)$$

From (2.2), the overall total cost of achieving the expected output quality vector $q_j = (q_{0j},...,q_{Kj})$ for m_j students in Department j when the number of research-active members of academic staff in Department j is n_j, is therefore:

$$C_j(q_j, m_j, n_j) = n_j\hat{C}_{0j} + m_j \sum_{h=1}^{K} \hat{C}_{hj} \;\; where\ \hat{C}_{hj} \equiv c_{hj}\chi_{hj} \;\; and\ c_{hj} \equiv (\omega_{1hj}/\alpha_h)^{\alpha_h}(\omega_2/(1-\alpha_h))^{1-\alpha_h}/A_h \quad (2.12)$$

Associated with any given level $\bar{C}_j$ of the overall total cost C_j in (2.12) we may define the *production possibility frontier (PPF)* given by:

$$q_{oj} = \varphi(q_{1j},...,q_{Kj}; \bar{C}_j, m_j, n_j) = \max_{r_{kj}} q_{0j} \;\; s.t.\ \text{(2.11), (2.12) \&} \; C_j(q_j, m_j, n_j) = \bar{C}_j \quad (2.13)$$

Under a university funding regime similar to that of HEFCE's (2010a), with its QR element of research funding, the income which any Department j receives will itself, however, increase with the value of its assessed *q*uality-weighted volume of research, as measured by $n_j q_{0j}$. In addition, the research grant income, G_j, that Department j receives from other sources may depend upon its number, n_j, of research-active members of academic staff and upon the average assessed quality of their research, as reflected in q_{0j}, so that $G_j = G(n_j, q_{0j})$. The funding for teaching that Department j receives under the type of funding regime we are considering here includes a positive monetary amount, η_1, for each of the specified number, m_j, of students which the central government funding agency agrees to fund, and for whom teaching quality assessments are required. However, we can also allow here for the possibility that Department j receives additional net income, H_j, from overseas postgraduate students who are not included in this central government-funded total, and who fall outside the teaching quality assessment regime we are considering here. Department j's ability to attract, and charge higher fees to, such students is assumed here to depend upon its academic reputation, as reflected in its research assessment q_{0j}. It may also depend upon the Department's teaching assessments, such that $H_j = H(q_j)$. Even if the teaching assessments do not relate directly to postgraduate teaching, as has been the case of the NSS for the period we are considering , the assessments are assumed to be widely available and used in the construction of published university league tables to which prospective postgraduate students may pay attention. Department j's overall net income will therefore equal:

$$Z_j = \eta_1 m_j + \eta_2 n_j q_{oj} + G(n_j, q_{0j}) + H(q_j) \equiv Z(q_j, m_j, n_j) \tag{2.14}$$

where $\eta_2 > 0$ is the central government research funding per *q*uality-weighted volume of research produced.

Since the assessed research quality, q_{0j}, of Department j is assumed to influence its academic standing in the eyes of academics, the wage ω_{1kj} which it must pay to attract and retain the time inputs of academics of any given level of ability is also assumed to depend upon q_{0j}, implying wage functions of the form:

$$\omega_{1kj} = \omega_{1k}(q_{0j}) \quad for\ k = 0,...,K \tag{2.15}$$

per unit of ability-weighted time input x_{kj}, but with the price ω_2 of supporting resources assumed to be independent of q_{0j}. In particular, we will assume that *ceteris paribus* research-active academic staff of a given level of research ability are willing to accept a lower wage per unit of their ability-weighted time input to work in a Department that is more prestigious according to its q_{0j} research rating. If this (marginal) effect becomes greater the greater q_{0j} becomes, we have:

$$\omega'_{10} \equiv (\partial\omega_{10} / \partial q_{0j}) / \omega_{10} < 0 \ \ and\ \omega''_{10} \equiv (\partial^2\omega_{10} / \partial q_{0j}^2) / \omega_{10} < 0 \tag{2.16}$$

For the sake of concreteness, we will assume that the research grant income function of Department j in (2.14) is of the form:

$$G(n_j, q_{0j}) = \kappa(q_{oj})n_j - \beta n_j^2 \ with\ \ \partial G / \partial n_j = \kappa(q_{oj}) - 2\beta n_j,\ \beta > 0\ and\ \partial\kappa(q_{oj}) / \partial q_{oj} > 0 \tag{2.17}$$

so that the additional income which an additional research-active member of academic staff generates for Department j is an increasing function of the assessed quality of its research output, but declines with n_j when there is some degree of spreading of the available research funds across Departments of different sizes. If the number of research-active academic staff which Department j hires is chosen to maximise its resource income net of its resource costs, we have from (2.11), (2.12), (2.14), (2.15) and (2.17):

$$n_j = n_j(q_j) \equiv [\kappa(q_{0j}) + \eta_2 q_{0j} - c_o(q_{0j}) \sum_{h=0}^{K} \tau_{0h} \upsilon_h(q_{hj})] / 2\beta \tag{2.18}$$

for $n_j(q_j) > 0$, with zero research-active staff hired if $n_j(q_j) \le 0$, and where

$$c_h(q_{0j}) = (\omega_{1h}(q_{0j}) / \alpha_h)^{\alpha_h} (\omega_2 / (1-\alpha_h))^{1-\alpha_h} / A_h \quad for\ h = 0,...,K \tag{2.19}$$

In order to identify an outer frontier of expected achievable performance for the assessed quality of Departmental research and teaching, we can now extend the inter-relationships that we include beyond those which are considered by the PPF in (2.13) to include the endogenous net income relationships (2.14) and (2.17) – (2.18), and the wage functions in (2.15). This yields an *expected achievement frontier (EAF)* defined by:

$$q_{oj} = \psi(q_{1j},...,q_{Kj}; m_j) = \max_{r_{kj}} q_{oj} \ s.t. \text{ (2.11), (2.12), (2.14) - (2.15), (2.17) - (2.18) and } C_j \le Z_j \tag{2.20}$$

which specifies the achievable combinations of expected research and teaching quality assessments which are feasible for any given value of the exogenously determined number of central government funded students, m_j, when Department j optimises the resourcing of its different activities and its number of research-active academic staff n_j within a net-income constraint that itself depends upon its assessed academic quality.

3. Non-Convexity and the Expected Achievement Frontier

We can combine the constraints involved in (2.20) into an overall net resource constraint given by:

$$\Phi(q_j, m_j) \equiv Z(q_j, m_j, n_j(q_j)) - \sum_{k=0}^{K} [\tau_{0k} n_j(q_j) c_0(q_{0j}) + m_j \sum_{h=1}^{K} \tau_{hk} c_h(q_{0j})] \upsilon_k(q_{kj}) = 0 \tag{3.1}$$

with (3.1) defining the achievable level of q_{0j} as an implicit function of $q_{1j},...,q_{Kj}$ and m_j. From (2.11) – (2.12), (2.14) – (2.18) and (3.1), we have:

$$\Phi_k = H_k - \varpi_{jk} \upsilon'_k \quad for\ k = 1,...,K \quad where\ \ \varpi_{jk} \equiv \tau_{0k} n_j c_0 + m_j \sum_{h=1}^{K} \tau_{hk} c_h \quad for\ k = 0,...,K \tag{3.2}$$

$$\Phi_0 = \eta_2 n_j + \kappa' n_j + H_0 - \varpi_{j0} \upsilon'_0 - \sum_{k=0}^{K} \varpi'_{jk} \upsilon_k \quad where\ \ \varpi'_{jk} \equiv \tau_{0k} n_j c'_0 + m_j \sum_{h=1}^{K} \tau_{hk} c'_h \quad for\ k = 0,...,K \tag{3.3}$$

where $\Phi_k \equiv \partial\Phi / \partial q_{kj}, \kappa' \equiv \partial\kappa / \partial q_{0j}, H_k \equiv \partial H / \partial q_{kj}, \upsilon'_k \equiv \partial \upsilon_k / \partial q_{kj} > 0, \omega'_{1k} \equiv (\partial \omega_{1k} / \partial q_{0j}) / \omega_{1k}, c'_k \equiv \omega'_{1k} \alpha_k c_k.$

Differentiation of (3.2) and (3.3), using (2.15) – (2.19), yields:

$$\Phi_{kk} = H_{kk} - \varpi_{jk} \upsilon''_k - (\tau_{0k} c_0 \upsilon'_k)^2 / 2\beta \quad for\ k = 1,...,K \tag{3.4}$$

$$\Phi_{00} = 2\beta(\partial n_j / \partial q_{0j})^2 + \kappa'' n_j + H_{00} - \varpi_{j0} \upsilon''_0 - 2\varpi'_{j0} \upsilon'_0 - \sum_{k=1}^{K} \varpi'_{jk} \upsilon'_k - \sum_{k=0}^{K} \varpi''_{jk} \upsilon_k \tag{3.5}$$

$$\Phi_{k0} = \Phi_{0k} = H_{0k} - \varpi'_{jk} \upsilon'_k - \tau_{0k} c_0 (\partial n_j / \partial q_{0j}) \quad for\ k = 1,...,K \tag{3.6}$$

where $$\Phi_{kh} \equiv \partial \Phi_k / \partial q_{hj}, \kappa'' \equiv \partial \kappa' / \partial q_{0j}, H_{kh} \equiv \partial H_k / \partial q_{hj}, \upsilon''_k \equiv \partial \upsilon'_k / \partial q_{kj}, \omega''_{1h} \equiv (\partial^2 \omega_{1h} / \partial q_{0j}^2) / \omega_{1h} \quad (3.7)$$

and $$\varpi''_{jk} \equiv \tau_{0k} n_j c''_0 + m_j \sum_{h=1}^{K} \tau_{hk} c''_h \;\; and \;\; c''_h \equiv \partial c'_h / \partial q_{j0} = \alpha_h c_h (\omega''_{1h} - (\omega'_{1h})^2 (1-\alpha_h)) \;\; for \; h = 0,...,K \quad (3.8)$$

The expected achievement possibility set, S, can be defined as the set of all (q_j, m_j) vectors satisfying the constraint:

$$\Phi(q_j, m_j) \geq 0 \quad (3.9)$$

in which the Department's costs of its expected achievements are no greater than its associated income. Convexity of S implies that Φ is quasi-concave, which in turn requires (see Arrow and Enthoven, 1961) that the associated bordered Hessian determinants weakly alternate in sign, with:

$$D_{0k} \equiv \begin{vmatrix} 0 & \Phi_0 & \Phi_k \\ \Phi_0 & \Phi_{00} & \Phi_{0k} \\ \Phi_k & \Phi_{k0} & \Phi_{kk} \end{vmatrix} = -[\Phi_k^2 \Phi_{00} + \Phi_0^2 \Phi_{kk} - 2\Phi_0 \Phi_k \Phi_{0k}] \geq 0 \quad (3.10)$$

for any $k = 1,...,K$. However, the necessary condition (3.10) for convexity of S is by no means guaranteed here, even for the simple case where we have:

$$0 = (\partial n_j / \partial q_{0j}) = \kappa'' = H_{kk} = H_{0k} = \omega'_{1h} \quad for \;\; h,k = 0,...,K \; and \; b_{kh} = 0 \;\; for \;\; h \neq k = 0,...,K \quad (3.11)$$

From (2.10), we have for $k = 0,...,K$:

$$\upsilon'_k = 1/ f'_k > 0 \; and \; \upsilon''_k = -f''_k / (f'_k)^2 \;\; where \;\; f''_k \equiv \partial f'_k / \partial v_{kj} = \sum_{g_k=1}^{G_k} \Delta w_k(g_k) \Pi_{kj} / \sigma_k^2 \quad (3.12)$$

$$and \;\; \Pi_{kj} \equiv \theta_{kj} \phi(\theta_{kj}) = -\partial \phi(\theta_{kj}) / \partial \theta_{kj} \;\; (>,=,<) \; 0 \; as \; v_{kj} (<,=,>) \gamma_{k,g_k} \;\; for \;\; \theta_{kj} \equiv (\gamma_{k,g_k} - v_{kj}) / \sigma_k \quad (3.13)$$

From (3.12) – (3.13) and (3.4) - (3.5), the condition $f''_k > 0$, and hence $\upsilon''_k < 0$ and $\Phi_{kk} > 0$ will be satisfied here if:

$$v_{kj} < \gamma_{k,1} \quad (3.14)$$

given (2.5), (2.7) and (2.10). However, (3.14) is unnecessarily strong for $\upsilon''_k < 0$, and hence $\Phi_{kk} > 0$ given (3.11), to hold. Rather from (3.12), we require simply that:

$$\sum_{g=g'_k+1}^{\hat{g}_k} \Delta w_k(g) \Pi_{kj}(\gamma_{k,g}) > \sum_{g_k=1}^{g'_k} \Delta w_k(g_k)(-\Pi_{kj}(\gamma_{k,g_k})) \; when \; \gamma_{k,g'_k} \leq v_{kj} < \gamma_{k,g'_k+1} \quad (3.15)$$

In the case of the 2008 RAE, for example, under the weighting system that is implicit in the associated QR funding element adopted by HEFCE (2010a) noted above, we have

$$w_0(4*) = 9, w_0(3*) = 3, w_0(2*) = 1, w_0(1*) = 0, and \; hence \; \Delta w_0(4*) = 6, \Delta w_0(3*) = 2, \Delta w_0(2*) = 1 \quad (3.16)$$

If the value of v_{0j}, for example, falls between γ_3 and γ_4, we have from (2.8) and (3.15) $\Pi_{0j}((\gamma_{02}-v_{0j})/\sigma_0)<0$ *and* $\Pi_{0j}(\gamma_{03}-v_{0j})/\sigma_0)<0$. (3.15) and (3.16) then require for $k=0$:

$$6\Pi_{0j}(\gamma_{04}-v_{0j})/\sigma_0)>2(-\Pi_{0j}(\gamma_{03}-v_{0j})/\sigma_0))+1(-\Pi_{0j}(\gamma_{02}-v_{0j})/\sigma_0))>0 \quad (3.17)$$

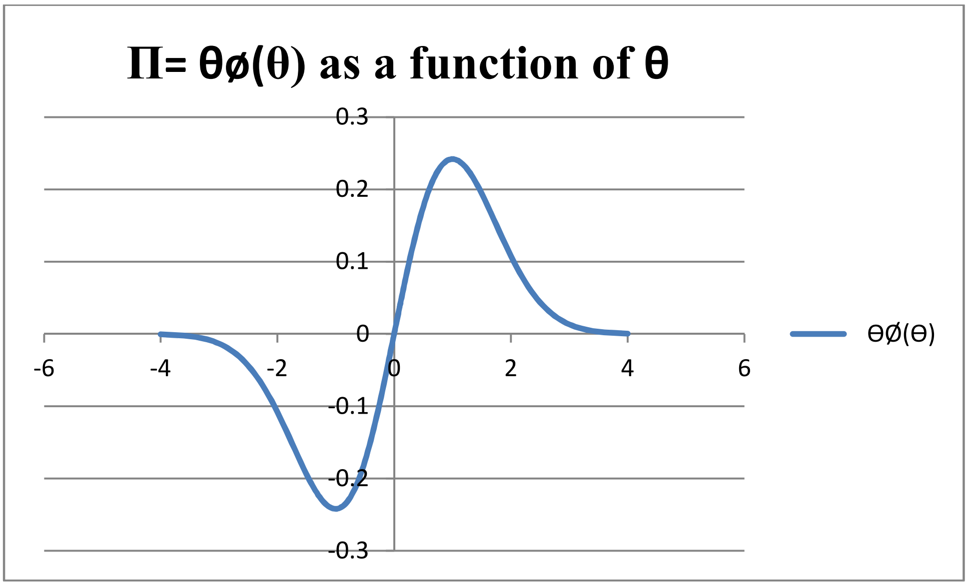

FIGURE 1

The maximum value which $-\Pi_{kj}$ can achieve can readily be shown to equal 0.2420, which occurs when $\theta_{kj}=-1$. It is then sufficient for (3.13), (3.15) and hence $\upsilon_0''<0$ to hold that:

$$6\Pi_{0j}(\gamma_{04}-v_{0j})/\sigma_0)>3(0.2420),\ \textit{and hence}\ 1.9214\sigma_0\geq\gamma_4-v_{0j}\geq0.3192\sigma_0 \quad (3.18)$$

using (3.12) and the numerical values of $\Pi=\theta\phi(\theta)$ as a function of θ. There is therefore a non-trivial range of the value of the resource inputs compared to the relevant performance hurdle which can result in $\upsilon_k''<0$ and hence $f_k''>0$, i.e. increasing marginal productivity of these resource inputs. Whenever $\Delta w_k(g_k)>0$, from (3.12) and (3.13), we also have:

$$\partial f_k''/\partial\gamma_{k,g_k}=[\Delta w_k(g_k)(1-\theta_{kj}^2)\phi(\theta_{kj})/(n_j\sigma_k^3)]>0\ \textit{iff}\ \ \gamma_{k,g_k}-\sigma_k<v_{kj}<\gamma_{k,g_k}+\sigma_k \quad (3.19)$$

so that a marginal raising of the hurdle level of performance in direction k for achieving a grade g_k will increase the rate of increase in the marginal productivity of more resources devoted to achievement by Department j in direction k, so long as the value v_{kj} of these resources is not more than one standard deviation σ_k away from the grade hurdle γ_{k,g_k}.

As Figure 1 illustrates, Π_{kj} in (3.12) – (3.13) is a highly non-linear function of θ_{kj}, and hence of the resourcing level v_{kj}. As well as there existing non-trivial ranges of the value of the resource inputs compared to the relevant performance hurdle which can result in $f_k''>0$, i.e. increasing marginal productivity of these resource inputs, there may therefore also be non-trivial ranges of the value of the resource inputs which result in $f_0'' < 0$, i.e. decreasing marginal productivity of these resource inputs over these ranges. SumR in both Figures 2 and 3 graphs the value of f_0'' in (3.12) for the case where the values of Δw_0 are given by (3.16) and we set $\sigma_0 = 1, \gamma_{0,2} = 0.5, \gamma_{0,3} = 3$, and $\gamma_{0,4} = 7$ over a range values for v_{0j} from 0 to 8. As can be seen from Figures 2 and 3, the associated value of f_0'' may change sign more than once over such a range.

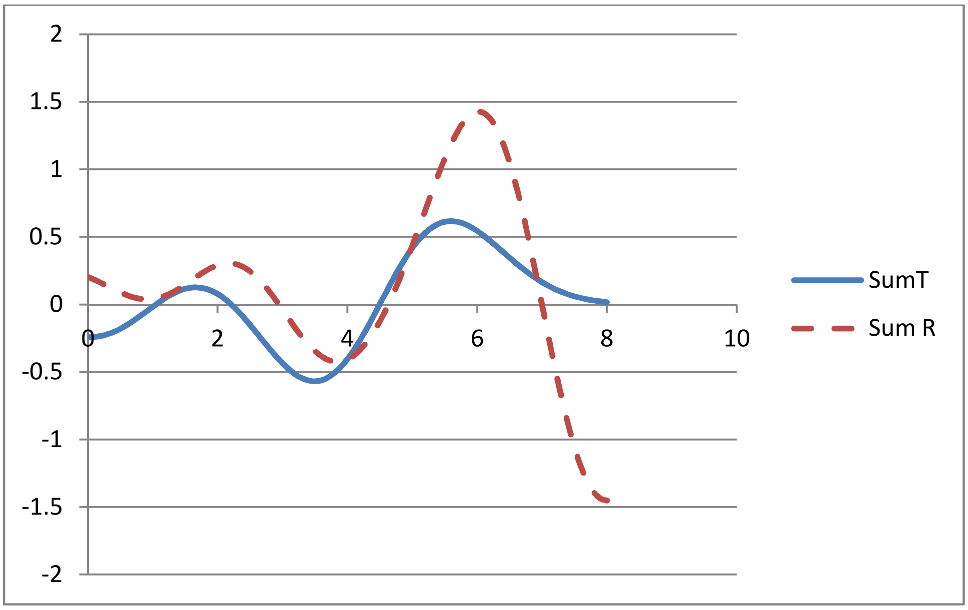

FIGURE 2

That there may be regions in which we have, at the same time, increasing marginal productivity for both research and teaching quality are also illustrated in Figures 2 and 3. The extent to which this occurs will depend *inter alia* upon the Δw_k weights that are placed upon the grade improvements that are assessed by the teaching quality exercise. Figure 2 illustrates the case where we set $\Delta w_1(2) = \Delta w_1(3) = \Delta w_1(4) = \Delta w_1(5) = 1$, and $\sigma_1 = 1, \gamma_{1,2} = 3, \gamma_{1,3} = 3.5, \gamma_{1,4} = 4, \gamma_{1,5} = 7$, with the total value of the resources devoted to the research and teaching activities $k = 0,1$ set equal to 8. There is here more than one region here is which both SumT, which shows the resultant value of f_1'' from (3.12), and SumR, which shows the value of f_0'', are positive, but also a region in which they are both negative. Figure 3 illustrates the case where we instead set $\sigma_1 = 1, \gamma_{1,2} = 0.5, \gamma_{1,3} = 1, \gamma_{1,4} = 3, \gamma_{1,5} = 4$, and

$$w_k(1) = w_k(2) = w_k(3) = 0, w_k(4) = w_k(5) = 1 \textit{ and hence } \Delta w_k(2) = \Delta w_k(3) = \Delta w_k(5) = 0, \Delta w_k(4) = 1 \quad (3.20)$$

for each teaching quality dimension $k = 1,...,K$, so that all student responses of grade 3 or below are treated as being "not satisfied" and those of grades 4 and 5 are treated as being "satisfied", as in HEFCE (2010b) . There is now one region in which both f_0'' and f_1'' are positive, and one region in which they are both negative. Under condition (3.11), the convexity condition (3.10) will then switch from being broken locally in the first region of the frontier to being satisfied locally over the second region. In addition, with SumT being close to zero and Sum R being positive in Figure 3 for many low values of the resources, v_{0j}, devoted to research, (3.4) – (3.7) and (3.11) may imply that the convexity condition (3.10) is broken also for many of these low values of v_{0j}.

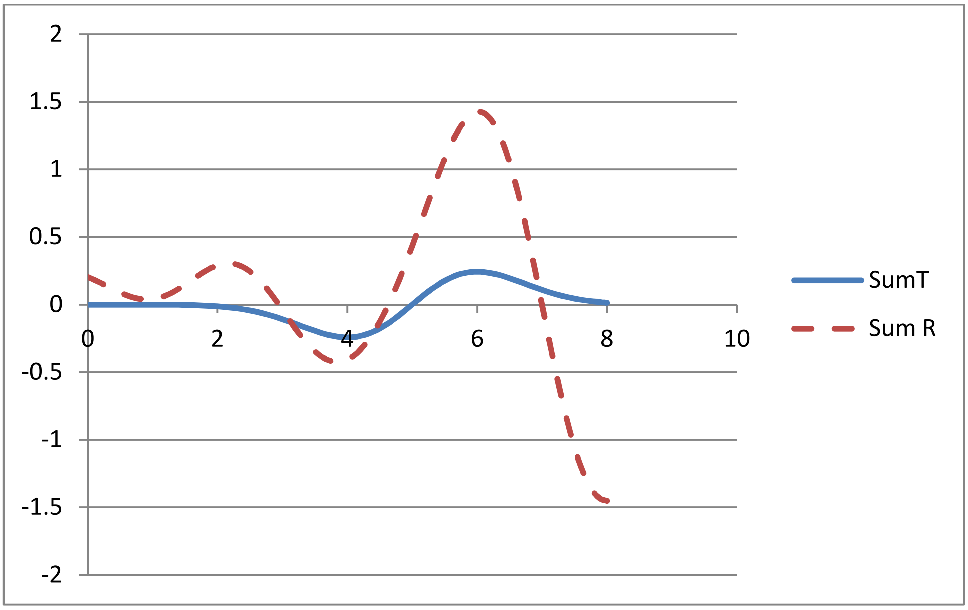

FIGURE 3

When we relax condition (3.11), there are *several other important sources of potential non-convexity* of the expected achievement possibility set. The first of these are that both κ'' and H_{00} may prove to be positive in (3.5). This may in particular be true if Department j's competitive success in securing research grant funding and additional net income from overseas postgraduate student fees, as reflected in κ in (2.17) and H in (2.14), are decreasing functions of Department j's *ranking* $R(q_{0j})$ from the top of a cumulative national distribution $F(q_{0j})$ of its research assessment across a total of s Departments nationally. We will then have:

$$\kappa(q_{0j}) = \widehat{\kappa}(R(q_{0j}))\ where\ R(q_{0j}) \equiv s[1-F(q_{0j})],\ \partial\widehat{\kappa}/\partial R < 0\ and\ \partial R/\partial q_{0j} = -s\vartheta(q_{0j}) \qquad (3.21)$$

where $\vartheta(q_{0j})$ is the density function associated with $F(q_{0j})$. We then have for $\vartheta'(q_{0j}) \equiv (\partial\vartheta(q_{0j})/\partial q_{0j})$:

$$\kappa' = -(\partial\widehat{\kappa}/\partial R)s\vartheta(q_{0j}) > 0,\ \kappa'' = (\partial^2\widehat{\kappa}/\partial R^2)(s\vartheta(q_{0j}))^2 - (\partial\widehat{\kappa}/\partial R)s\vartheta'(q_{0j}) \qquad (3.22)$$

with a similar equation holding also for H_{00}. If Department j's competitive position increases significantly the closer it gets to the top of the national distribution, we may have $\partial^2\widehat{\kappa}/\partial R^2$ and $\partial^2 H/\partial R^2$ both positive, which will yield overall positive values to κ'' and H_{00} in (3.5) whenever $\vartheta'(q_{0j}) \geq 0$ or whenever $-\vartheta'(q_{0j})$ is sufficiently small. If $\vartheta(q_{0j})$ is unimodal, κ'' and H_{00} will also have larger positive values the closer q_{0j} is to the mode of this distribution, reflecting *the large change in the national ranking of Department j which can be achieved by a relatively small change in its research assessment* q_{0j} when the Department is close to the mode of the national distribution. Similar reasoning will imply $H_{kk} > 0$ in (3.4) if q_{kj} impacts upon H_j in a parallel way via the national distribution for q_{kj}, which will in turn imply $\Phi_{kk} > 0$ in (3.4) whenever $\upsilon_k'' < 0$ and B is a diagonal matrix.

Moreover, if an improved research standing impacts upon the wage function in (2.15) via its ranking $R(q_{0j})$, so that the further Department j is from the top of the national distribution the higher the wage premium it must pay to attract and retain staff of a given level of ability in direction h, we will have:

$$\omega_{1h}(q_{0j}) = \hat{\omega}_{1h}(R(q_{0j})), \partial\hat{\omega}_{1h} / \partial R > 0, \omega_{1h}' < 0, \omega_{1h}'' = [(\partial^2 \hat{w}_{1h} / \partial R^2)(s\vartheta(q_{0j}))^2 - (\partial\hat{\omega}_{1h} / \partial R)s\vartheta'(q_{0j})] / \omega_{1h} \quad (3.23)$$

If the prestige of working in the Department increases significantly as it gets closer and closer to the top of the national distribution, we may have $\partial^2\hat{\omega}_{1h} / \partial R^2 < 0$, with a greater weight on this term that contributes to a larger negative value of ω_{1h}'' in (3.7) - (3.8) when q_{0j} is close to the mode of national distribution. In addition, we will have $\vartheta'(q_{0j}) \geq 0$ for all values of q_{0j} less than or equal to the mode, confirming the overall negative value to ω_{1h}'' in (3.23) for these values to q_{0j}. Such negative values to ω_{1h}' and ω_{1h}'' in (3.23), and hence also of ϖ_{jk}'' in (3.8) whenever $\alpha_k \leq 1$, will reinforce the positive components to Φ_{00} in (3.5) when B is a diagonal matrix. A similar conclusion indeed applies to any non-zero value to $(\partial n_j / \partial q_{0j})$ in (3.5) and (2.18). Such positive components will tend to increase the degree of non-convexity in regions in which both f_0'' and f_1'' are positive, and reduce the extent to which the local convexity requirement (3.10) holds elsewhere.

However, if the wage effect ω_{1k}' for $k \neq 0$ in (3.23) and (3.3) and the resultant value to ϖ_{jk}' in (3.3) are strongly negative, these positive components to Φ_{00} in (3.5) may be offset to some extent by positive elements to Φ_{0k} in (3.6) and (3.10). A strong wage effect $\omega_{1k}' < 0$ for $k \neq 0$ will imply that boosting the Department's research quality significantly reduces the cost of it attracting able teachers, thereby generating a pecuniary form of positive economy of scope that in itself reinforces the degree of local convexity of the expected achievement possibility set.

When we relax the assumption that the synergy matrix B is diagonal, the scope for less *a priori* predictability of the extent of the overall departure from convexity of the expected achievement possibility set is increased. In particular, negative off-diagonal elements of the matrix inverse $\Gamma \equiv [\tau_{hk}] \equiv B^{-1}$ may be generated by positive synergies between the resources devoted to increasing research quality and those devoted to increasing some aspects of teaching quality, such as the extent to which teaching is intellectually stimulating in its content. These negative values of τ_{hk} for $h \neq k$ will provide partial offsets to reduce the extent to which Φ_{00} and Φ_{kk} are positive in (3.4) and (3.5). Similarly, a positive value of H_{0k} due to complementarity between research and teaching quality in the extent to which they increase the attractiveness of the Department to overseas students paying higher tuition fees will provide a partial offset in (3.6) and (3.8) to the extent to which the local convexity condition (3.10) is broken. Moreover even if $\tau_{0h} > 0$ for $h = 1,...,K$, it may be the case that a higher Departmental research performance attracts more research-orientated staff who are less willing to spend time on other labour-intensive teaching activities, such as marking large volumes of essays, so that $\omega_{1h}' > 0$ and $\omega_{1h}'' > 0$ in (3.5) and (3.8), for these other directions $h \neq 0$. Thus while there may indeed be sections of the EAF for which the convexity condition (3.10) is broken, its universal breach, in the form of global strict quasi-convexity of the function Φ in (3.1) that defines the EAF, is also not guaranteed. As in Figure 4, there may then be sections of the EAF that are locally concave from above and other sections which are locally convex from above. Different Departments may be in equilibrium at different points, such as W, A, L, M and D along the EAF. This will depend *inter alia* upon their willingness in their resource allocation decisions to make trade-offs between their assessed research quality and their assessed teaching quality. If the preferences of individual Departments between research and teaching quality are representable by utility functions $U_j(q_j)$ for $j = 1,...,s$, the second-order conditions for a constrained maximum to $U_j(q_j)$ subject to the constraint (3.7) will be satisfied if the associated indifference curves have a sufficiently large curvature in each relevant direction, such as at points A, L or M in Figure 4, as dependent upon the position of each individual Department's indifference curves between their research and teaching quality scores. The existence of sections of the EAF which are locally strictly concave from above will, however, tend to make increases in teaching quality more costly to achieve along at least part of the frontier in terms of the accompanying fall in research quality.

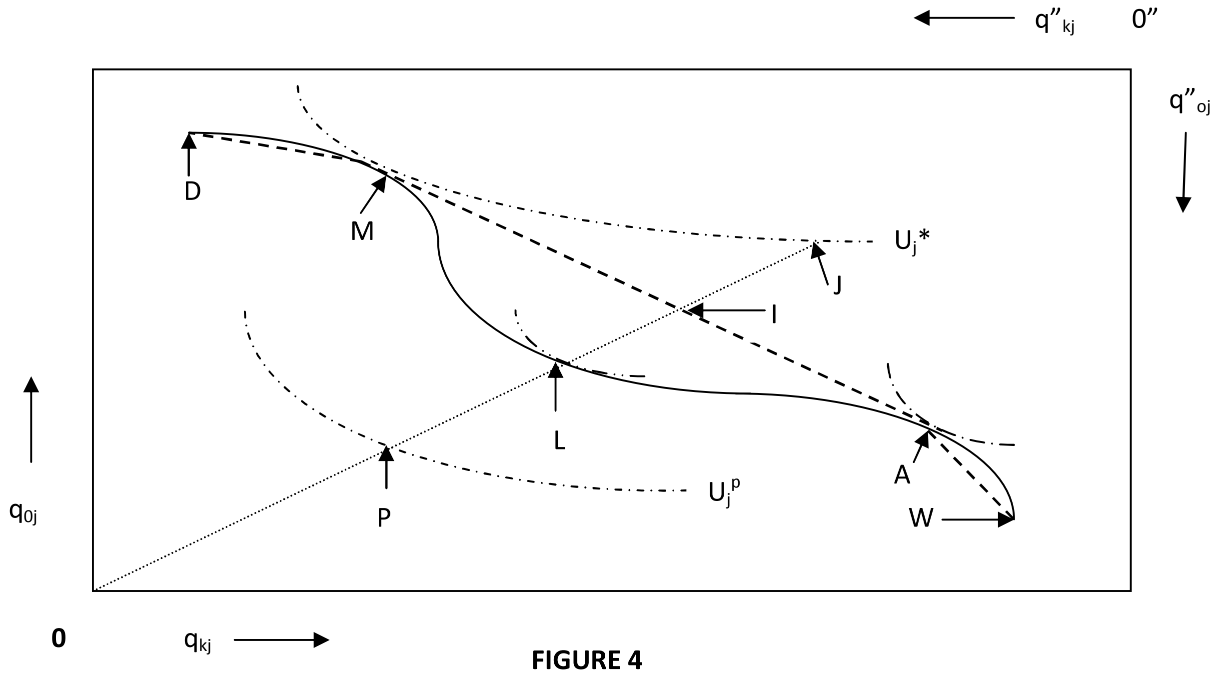

FIGURE 4

4. FDH Analysis and DEA

Once constraints on the available resources imply potential trade-offs between the time and other resources devoted to research and to teaching activities, it becomes important to also examine empirically the nature of the achievement possibility frontier between these different assessed outputs and the extent to which some Departments fall inside this frontier. One approach to estimating the position of such a frontier is that of stochastic frontier analysis (SFA) (see Kumbhakar and Lovell, 2000), which seeks to estimate a particular functional form of a parametric production or cost function, that is subject to additive disturbance terms that are assumed to be decomposable into a normally distributed term reflecting ability or other features of the resource endowment of the producer, such as the vintage of their capital stock, and a half-normally distributed non-positive disturbance term reflecting the efficiency of the producer. However, difficulties arise in this context of finding a suitable parametric functional form that permits more than one switch between strictly convex and strictly concave sections of the frontier. It is notable that the coefficients of the multi-output CES cost function which Izadi et al (2002) estimate for the quality-unadjusted volumes of output of higher education institutions in the UK using SFA imply that the resultant cost function is strictly quasi-concave for all positive values of the output (see also Baumol et al, 1982, p. 461), with an associated non-convex feasible set of outputs for any given level of total expenditure . Moreover, once the assessed quality of the output becomes our main focus of interest and funding is partially endogenous, such a single-equation estimation of a cost function through SFA becomes less appropriate. Furthermore, while published cost data are readily available for the UK at institutional level, the use of such aggregate institutional data "often can be misleading" according to Dunbar and Lewis (1995, p. 120) because of "quite dissimilar production functions" across different academic disciplines, such as Chemistry and English. The lack of published detailed Departmental cost data for individual UK Departments of Economics presents an additional substantial impediment to such an estimation in our present context. Similarly, the lack of reliable detailed data on the resources devoted to each separate direction of achievement by each such UK Department impedes the direct estimation of the underlying production functions for their assessed output quality in (2.3) – (2.5) using techniques such as ordered probit analysis (see Cameron and Trivedi, 2005).

One main non-parametric approach which has been widely used (see e.g. Emrouznejad et al, 2008) for assessing the efficiency and productivity of multi-output firms, and which has been applied also in the schools sector of education (see e.g. Jesson, Mayston and Smith, 1987; Mayston and Jesson, 1988), and suggested for use also in the higher education sector (Johnes, 1999), is that of Data Envelopment Analysis (DEA). When the focus of interest is on the extent to which individual academic Departments are capable of feasible improvements in their assessed outputs, particularly relevant here is the *output-orientated* version of the DEA model due to Banker, Charnes and Cooper (1984) (BCC) , which is formulated in terms of the linear program:

$$\max \varsigma_j \ s.t. \ X\lambda \le X_j, \varsigma_j Y_j - Y\lambda \le 0, e\lambda = 1, \lambda \equiv (\lambda_1, ..., \lambda_n)' \ge 0, e \equiv (1,...,1) \quad (4.1)$$

and which relaxes the earlier formulation of DEA by Charnes, Cooper and Rhodes (1978) (CCR) in which constant returns to scale are assumed (see Cooper, Seiford and Tone, 2007). The DEA model (4.1), in which X_j and Y_j are the input and output vectors, respectively, of Department j, with $X \equiv (X_1, ..., X_s)$ and $Y \equiv (Y_1, ..., Y_s)$, involves finding the maximum radial expansion in the output vector of Department j that will still be feasible within the estimated possibility set formed by taking convex combinations of the input and output vectors of Departments in the observed set. The reciprocal of the estimated ς_j coefficient is then taken to be a measure of the effectiveness, or 'technical efficiency', of the Department's actual output vector compared to what DEA takes to be feasible based upon the frontier formed by the convex hull of all Department's performances. However, a major problem with the use of DEA in our current context arises when its underlying assumption of convexity of the feasible set does not hold, but instead the actual achievement frontier is strictly concave from above over some significant region. DEA's comparison with projected points on this convex hull then involves a comparison with input-output vectors which are in fact infeasible. This is illustrated in Figure 4, where WALMD represents the actual achievement frontier for the assessed outputs for a given level of input, with a non-convex region to the associated feasible set between points A, L and M. DEA would here estimate an efficient point such as L to have a level of technical efficiency of only (OL/OI) < 1 based upon its comparison with the infeasible point I that lies on the artificial frontier formed from a convex combination of the adjacent points A and M. An inefficient point, such as P in Figure 4, that lies inside the feasible frontier WALMD, would similarly be given an excessively low efficiency score of only OP/OI in Figure 4, and implied excessive ability to make proportional improvements of OI/OP in each of its outputs, compared to the proportional improvements of OL/OP that are actually feasible, with an associated true 'technical efficiency' score of OP/OL > OP/OI.

Regions of non-convexity of the feasible set also have important implications for a Department's *price- or allocative-efficiency*, as a reflection of the extent to which a re-balancing of the proportions in which the different outputs are produced may be desirable. Given the non-convex feasible set associated with the frontier WALMD in Figure 4, the optimal overall position for a Department j with a utility function $U_j(q_j)$ between the relevant dimensions of achievement that has the associated indifference curves passing through points P and M in Figure 4 would be at point M. If the actual position achieved by the Department is at P, an overall value-based measure of its effectiveness would be given by $\xi_{jP} = U_j^P / U_j^*$, as dependent upon the values the Department places upon achievements in each dimension of research and teaching within its utility function. If $U_j(q_j)$ is homothetic, a consistent measure of ξ_{jP} is provided by the ratio OP/OJ, where J is the point of intersection of Department j's indifference curve with the ray OP through the origin at 0. ξ_{jP} may in turn then be expressed as the product of the technical measure of efficiency OP/OL and the value-based measure of allocative efficiency OL/OJ. Such a measure of allocative efficiency provides a generalisation of the concept of price efficiency introduced by Farrell (1957) for the case of linear iso-valuation curves when there is a constant price for each unit of output.

In contrast to the case of a convex feasible set considered by Farrell, the existence of non-convex regions of the feasible set may well result in values of the measure of allocative efficiency, such as OL/OJ, which are substantially less than one, and which are at least as great a source of the overall value-based measure of its effectiveness being low as is its technical inefficiency. The non-convex region of the feasible set in Figure 4 involves gains from increasing marginal productivity in individual directions along the frontier, with a greater potential allocative loss arising if Department j does not reap these gains from specialisation in the directions of achievement that offer the highest relative valuation according its utility function. Once one moves from a market-oriented producer facing constant market prices for its outputs to a not-for-profit producer, such as a university Department, producing outputs without a constant market price, knowledge of its utility function, and associated subjective trade-offs between achievements in each relevant direction, would be required for an assessment of its value-based allocative and overall efficiency. In this paper we therefore focus empirically upon assessing Departmental technical efficiency, but with the identification of possible regions of non-convexity of the feasible set important both for accurately assessing technical efficiency and for identifying regions where the issues raised by allocative efficiency may also be particularly acute.

As noted above, DEA does not provide reliable measures of technical efficiency when there are regions of non-convexity of the feasible set. An alternative non-parametric model for estimating the performance of Departments, and their scope for improvement, compared to what has been shown to be feasible by the input-output vectors of all Departments in the sample is provided by the Free Disposal Hull (FDH) model developed by Duprin et al (1984). In the context of our current focus on the assessed quality-adjusted output vector q_j of each Department $j \in \Xi$, and its associated exogenously determined input m_j, the FDH model seeks to identify the possibility set :

$$\Omega \equiv \{(m_j, q_{0j}, ..., q_{Kj}) \mid m_j \geq m_\iota, q_{0j} \leq q_{0\iota}, ..., q_{Kj} \leq q_{K\iota} \ \ for\ some\ \iota \in \Xi\} \qquad (4.2)$$

that, as in Figure 5, involves a step function of input-output points with at least as much input, and no more output, than the input-output vector that some existing Department $\iota \in \Xi$ has shown to be feasible. The original

FDH model of Duprin et al (1984), and the currently available software (DEA-solver) for estimating this model (see Cooper, Seiford and Tone, 2007), involve the *input-orientated* mixed integer programming formulation:

$$\min \rho_j \ s.t.\ \rho_j X_j - X\lambda \geq 0, Y_j - Y\lambda \leq 0, e\lambda = 1, \lambda_\iota \in \{0,1\} \ for\ each\ \iota \in \Xi \quad (4.3)$$

However, as noted above, an output-orientation is needed in our current context to show the extent of the feasible improvements in output quality for each Department for their given student intake, rather than the reduction in student numbers that would be consistent with their existing output quality vector. Fortunately, we may achieve such an output-orientated formulation of FDH analysis by considering the *performance shortfall* $q''_j \equiv (q^*_{oj} - q_{oj},...,q^*_{Kj} - q_{Kj})'$ of each Department j's existing output quality vector q_j compared to the maximum scores $q^*_j \equiv (q^*_{0j},...,q^*_{Kj})'$ that Department j could have attained under the quality assessment procedure. For the RAE this involves a maximum score of 9.0, and for each of the 13 NSS teaching quality questions it involves a maximum score of 1.0 (i.e. 100 per cent satisfaction) under the weighting systems (3.16) and (3.21), i.e.

$$q^*_j \equiv (q^*_{0j}, q^*_{1j},...,q^*_{13j})' = (9,1,...,1)' \ for\ each\ j \in \Xi \quad (4.4)$$

When we compare Department j's existing input of m_j students with some upper bound $\tilde{m}$ of all Departmental student numbers in the sample, we can then re-orientate (4.3) to one which estimates the extent of the output quality performance shortfall through solving:

$$\min \rho_j \ s.t.\ \rho_j(q^*_j - q_j) - (Q^* - Q)\lambda \geq 0, (\tilde{m} - m_j) - (\tilde{\tilde{m}} - m)\lambda \leq 0, e\lambda = 1, \lambda_\iota \in \{0,1\} \ for\ each\ \iota \in \Xi \quad (4.5)$$

where $Q^* \equiv (q^*_1,...,q^*_s), Q \equiv (q_1,...,q_s), m \equiv (m_1,...,m_s)$ and $\tilde{\tilde{m}} \equiv (\tilde{m},...,\tilde{m})$.

with the corresponding DEA model under the BCC formulation (see Cooper, Seiford and Tone, 2007, p. 91) solving the linear program:

$$\min \rho_j \ s.t.\ \rho_j(q^*_j - q_j) - (Q^* - Q)\lambda \geq 0, (\tilde{m} - m_j) - (\tilde{\tilde{m}} - m)\lambda \leq 0, e\lambda = 1, \lambda_\iota \geq 0 \ for\ each\ \iota \in \Xi \quad (4.6)$$

where the results in (4.5) and (4.6) are independent of our choice of $\tilde{m}$. As in Figure 5 (where for simplicity we hold the input constant at $\tilde{m}$), each Department's effectiveness is now assessed in terms of the extent to which its output quality shortfall can be reduced along a ray from the origin O' formed by the maximal output vector Q^*, whilst remaining on the achievement frontier for its given level of input. In contrast to the under-estimate which DEA generates of the effectiveness of efficient Departments along a concave-from-above section of the actual achievement frontier, the achievement frontier which the FDH model constructs will pass through all efficient points, such as point L in Figure 5, and hence give all such efficient Departments an effectiveness score of 100 per cent. The effectiveness score of Department P in Figure 5 will be assessed under FDH as $(O'H / O'P)$, with an additional output slack in direction k of HL that a comparison with the actual performance of Department L suggests could be remedied in addition to the overall proportional reduction of $(O'H / O'P)$ in Department P's performance shortfall in its output quality vector. Given the existence of such additional slacks, it is of interest to record not only the overall effectiveness score $(O'H / O'P)$ of each such Department P but also their effectiveness scores in each relevant dimension of their output after taking account of any such additional slacks. This can be assessed here as *the ratio of their projected quality shortfall to their actual quality shortfall*, along each relevant dimension of their output. Thus if point P in Figure 5 has coordinates (q''_{kP}, q''_{0P}) from the origin O', and the quality shortfall coordinates of the actual Department L with which P is being compared under FDH are (q''_{kL}, q''_{0L}), Department P's effectiveness score in teaching quality k is given by the ratio q''_{kL} / q''_{kP} and its effectiveness score in the research quality dimension 0 is given by q''_{0L} / q''_{0P}. For non-dominated observed points, such as D, M, L, A and W in Figure 5, on the efficiency frontier identified by FDH, their effectiveness scores in each dimension would be 1.0.

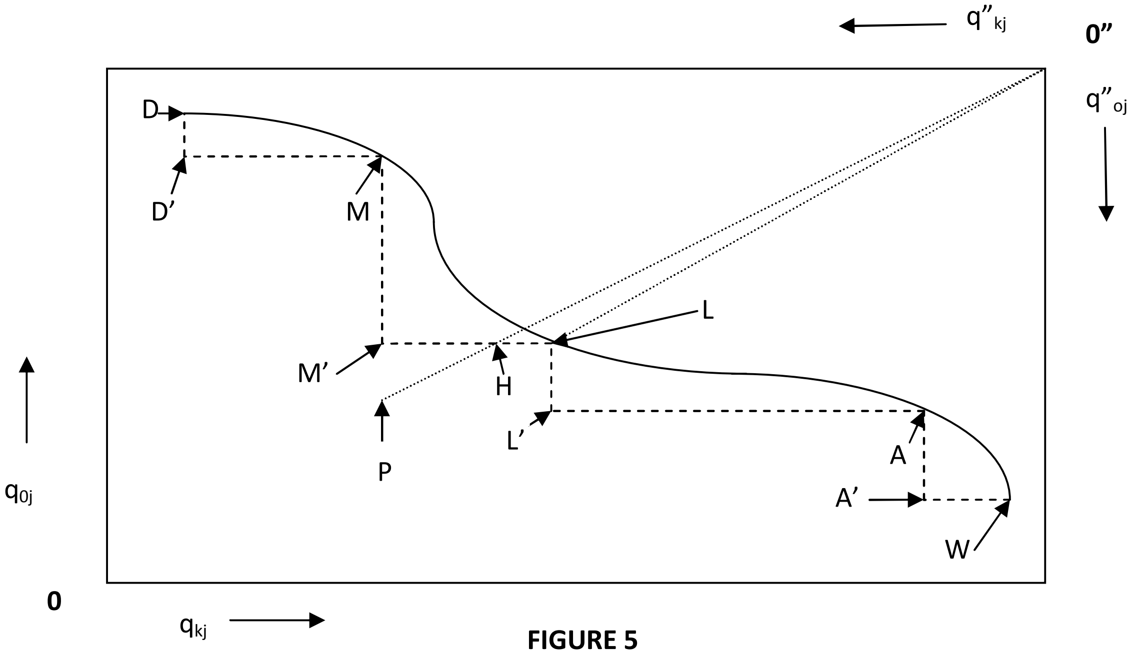

FIGURE 5

If the points D, M, L, A and W in Figure 5 all lay along an underlying efficiency frontier to a convex feasible set, we would not observe:

$$q''_T \equiv \sum_{i\in\Psi} \lambda_i (q_i^* - q_i) < (q_L^* - q_L) \; and \; m''_T \equiv \sum_{i\in\Psi} \lambda_i (\tilde{m} - m_i) \geq (\tilde{m} - m_L) \quad for\ some\ \lambda_i : \lambda_i \geq 0 \; and \; \sum_{i\in\Psi} \lambda_i = 1 \quad (4.7)$$

where Ψ is a set of points on the frontier not including L and the vector inequality $<$ involving here a strict inequality for at least one component, and at least weak inequalities for all components , of the relevant vectors. If (4.7) does hold, it implies that the convex combination at T dominates the point L on the frontier, contrary to the assumption of convexity of the feasible set. However, if there is a non-convex region to the feasible set with an efficiency frontier passing through points such as D, M, L, A and W in Figure 5, a convex combination at a point such as T, of the observed points A and M, may satisfy (4.7), though with T here falling outside the feasible set. Because of its focus on convex combinations of observed points, DEA may then report a point such as L as being inefficient, based upon a reference set of other observed points, such as A and M, involved in the convex combination, even though FDH estimates L, A and M to be efficient. In doing so, it would confirm that (4.7) holds for this reference set Ψ, and hence that any efficiency frontier passing through L, A and M does not satisfy the requirements of convexity over the region Λ_Ψ formed by convex combinations of the points in Ψ. Thus even if there is a smooth underlying efficiency frontier passing through points such as L, A and M in Figure 5, rather than FDH's step function, it will still not satisfy the convexity condition over this region.

5. The Empirical Results

In this section we illustrate empirically the extent of the differences which the choice of analytical technique can make to the effectiveness estimates of individual academic Departments in the context of research and teaching quality assessments, when we relax the strong assumption of convexity of the achievement possibility set, that is implicit in the use of DEA, to permit possible regions of non-convexity through the use of FDH analysis. Since research and international student fee incomes are treated as endogenous in the definition of the achievement possibility set, our choice of the exogenous variable which we will use to allow for possible differences in the externally imposed constraints on individual Departments will be their total home (i.e. UK) and other EU student numbers on undergraduate or taught Masters programmes. These numbers formed the basis for an externally driven net budget constraint for funding such students in individual institutions that was imposed by university funding agencies for the UK, such as HEFCE (2006) and the Scottish Funding Council (see Universities Scotland, 2008), during the period of our study. While individual universities still retained some freedom on how they allocated these totals across individual academic Departments, we will assume here that for any individual university i the total home and other EU student numbers which any given academic Department j was permitted to recruit was a fixed proportion π_{ij} of university i's total. Fortunately, data on the total home and other EU student numbers across their principal subjects of study are available from the UK's Higher Education Statistical Agency (HESA), even though for many individual academic Departments of interest, such as Economics Departments, there are no data published on individual departmental expenditure patterns.

In order to investigate empirically the achievement possibility set between assessed research and teaching quality, we will analyse the results of the latest available Research Assessment Exercise (RAE) for our main subject of interest in this paper, namely Economics, that was carried out in 2008 on the basis of submissions for research-active staff on the census date of 31st October 2007, alongside the published results for teaching quality assessments for Economics from the National Student Survey (NSS) in the academic year 2006-7. There is a total of 50 university Departments of Economics for which NSS results are available for this academic year, with 29 of these having positive entries for the corresponding RAE but with a zero research quality assessment for the remaining 21. Using the weighting systems (3.16) and (3.21), together with (4.4), we can compare the results of applying the FDH model (4.5) to the individual Departmental research and teaching quality assessments derived from their RAE results and their NSS scores for each of its 13 teaching quality questions, with those from the corresponding DEA model in (4.6).

Table 1 below shows the overall result of this comparison in terms of the arithmetic means of the effectiveness coefficients across all Departments in the sample under both FDH analysis and DEA, and the ratio between their respective effectiveness estimates. In addition to their estimates of an overall effectiveness coefficient ρ_j for each Department j in (4.5) and (4.6), both FDH analysis and DEA may generate slack variables, of the additional improvements which are possible in any given relevant dimension on top of the overall proportionate improvements which are implicit in their estimates of the overall effectiveness coefficient ρ_j for each Department. If these additional improvements are positive in any given direction, it will correspondingly further reduce the estimated effectiveness coefficient for this dimension of achievement below the overall effectiveness coefficient. Table 1 therefore includes the arithmetic means across the Economics Departments of the resultant effectiveness scores for each individual dimension of achievement that are estimated by both FDH analysis and DEA.

Overall, and for many individual dimensions of achievement, the mean value of the coefficient of effectiveness for Economics Departments estimated by the FDH model is higher in Table 1 by a factor of more than ten per cent times that estimated by DEA. For the NSS Q4. on whether the course was intellectually stimulating, the mean coefficient of effectiveness found under FDH analysis was higher by a factor of more than 25 per cent times that estimated by DEA. Across the 50 individual Departments, the effectiveness coefficients that are estimated by our FDH analysis and DEA are positively correlated, both according to their Pearson correlation coefficients and their Spearman rank correlation coefficients, as shown in columns (i) and (iv) in Table 2. However, all of the correlation coefficients in columns (i) and (iv) are less than 0.8, with the exception of the Spearman rank correlation coefficient between the overall effectiveness coefficients of individual Departments that are estimated by our FDH analysis and DEA, which is very close to 0.8. A total of 30 Departments (i.e. 60 per cent of the 50) were found to be fully effective under our FDH analysis, with an overall effectiveness score of 1.0 and zero slacks on each individual dimension of achievement. This is in contrast to the total of only 17 Departments (i.e. 34 per cent of the 50) which were estimated by DEA to be fully effective. All of those found to be fully effective under DEA were found to be fully effective under FDH analysis, though not *vice versa*. The choice of analytical technique for assessing effectiveness therefore makes a difference both to the rankings of individual Departments and to the estimated scope that exists for their individual improvement along each assessed dimension of achievement.

Columns (ii) and (v) of Table 2 show the Pearson and Spearman rank correlation coefficients across the 50 Departments between the effectiveness coefficients produced by FDH analysis for each individual assessed dimension of achievement and the raw data for these achievements by the individual Departments. These correlation coefficients are all less than 0.6, illustrating that making judgements based simply upon the individual raw data for these achievements, without taking account of the trade-offs which even fully effective Departments must make between the different dimensions of achievement along an efficiency frontier, will imperfectly reflect what is implied to be feasible by a technique, such as FDH analysis, which seeks to estimate the efficiency frontier. This conflict appears to be greatest in the case of the NSS Q.9 on whether feedback has been sufficient to help the student clarify things they did not understand, where the Pearson and Spearman rank correlation coefficients between the results of our FDH analysis and the individual raw data are – 0.458 and – 0.708 respectively. If giving feedback to individual students is labour intensive, it may indeed conflict with the resources which can be devoted to improving other dimensions of achievement, such as research quality. When adjustment is made for Departmental achievements in these other directions, the pattern of feasible scores for such feedback across individual Departments may therefore differ significantly from the individual score distribution within the unadjusted raw data.

Similar negative correlations for the NSS Q.9 are found between the raw data and the results of the DEA frontier analysis in columns (iii) and (vi) in Table 2. For all other dimensions of achievement, the correlation coefficients across individual Departments between the raw data scores and the results of DEA in columns (iii) and (vi) are higher than those produced by our FDH analysis in columns (ii) and (v) respectively. As noted above, FDH analysis, unlike DEA, permits regions of non-convexity of the feasible set, with associated increasing marginal trade-offs between the different dimensions of achievement along the efficiency frontier. *The results under FDH analysis will in general therefore tend to bring out more strongly the conflict which exists between making judgements of Departmental performance on the basis of the unadjusted raw data and those based upon an estimate of the* *feasible achievement frontier that takes into account these trade-offs between the different dimensions of achievement.*

Examination of the results of applying FDH analysis and DEA to the data set for individual Departments reveals that there are 13 regions where the non-convexity condition (4.7) holds. In each of these cases, there is a Department which FDH analysis finds to be fully effective but which DEA does not, with the Department which FDH analysis finds to be fully effective not entering into the reference set Ψ which DEA uses to assess its effectiveness score.

	FDH	DEA	FDH:DEA
Overall Effectiveness Coefficient	0.933	0.845	1.104
Research Assessment Exercise Effectiveness Coefficient	0.901	0.829	1.087
Teaching Effectiveness Coefficients on NSS Responses to:			
Q1. Staff are good at explaining things	0.800	0.745	1.073
Q2. Staff have made the subject interesting	0.768	0.666	1.153
Q3. Staff are enthusiastic about what they are teaching	0.714	0.619	1.153
Q4. The course is intellectually stimulating	0.804	0.634	1.269
Q5. The criteria used in marking have been clear in advance	0.760	0.723	1.051
Q6. Assessment arrangements and marking have been fair	0.810	0.707	1.145
Q7. Feedback on my work has been prompt	0.814	0.727	1.119
Q8. I have received detailed comments on my work	0.812	0.743	1.093
Q9. Feedback on my work has helped me clarify things I did not understand	0.863	0.764	1.130
Q10. I have received sufficient advice and support with my studies	0.808	0.740	1.092
Q11. I have been able to contact staff when I needed to	0.783	0.697	1.123
Q12. Good advice was available when I needed to make study choices	0.790	0.755	1.047
Q22. Overall, I am satisfied with the quality of the course	0.778	0.660	1.177

TABLE 1: Mean Values of the Departmental Effectiveness Scores

	Pearson Correlations			Spearman Rank Correlations		
EFFECTIVENESS SCORES	(i)FDH-DEA	(ii)FDH-Raw	(iii)DEA-Raw	(iv)FDH-DEA	(v)FDH-Raw	(vi)DEA-Raw
Overall	0.785	-	-	0.801	-	-
Research Assessment	0.772	0.347	0.451	0.798	0.288	0.396
NSS Q1.	0.766	0.507	0.677	0.743	0.487	0.656
NSS Q2.	0.745	0.281	0.502	0.745	0.355	0.457
NSS Q3.	0.783	0.342	0.579	0.759	0.469	0.551
NSS Q4.	0.601	0.253	0.447	0.629	0.284	0.527
NSS Q5.	0.685	0.265	0.333	0.652	0.326	0.383
NSS Q6.	0.625	0.447	0.671	0.643	0.478	0.704
NSS Q7.	0.726	0.570	0.705	0.736	0.570	0.686
NSS Q8.	0.745	0.535	0.720	0.753	0.528	0.699
NSS Q9.	0.620	-0.458	-0.708	0.648	-0.468	-0.738
NSS Q10.	0.742	0.479	0.624	0.748	0.509	0.607
NSS Q11.	0.717	0.315	0.529	0.701	0.225	0.462
NSS Q12.	0.791	0.394	0.498	0.775	0.381	0.478
NSS Q22.	0.676	0.462	0.784	0.666	0.507	0.781

TABLE 2: Correlation Coefficients Across individual Departments in Effectiveness Scores

6. Conclusion

With universities and their academic Departments under increasing pressure to improve the assessed quality of their research and teaching in the UK and elsewhere, the estimation of the feasible set of joint achievements which are possible in each relevant direction becomes of heightened importance. In estimating the associated achievement possibility frontier based upon what has been demonstrated to have been achievable by existing Departments, account must be taken of the dependence of the assessed quality of research and teaching upon the mapping of what may be underlying continuous variables into a limited number of discrete quality grades, via associated quality thresholds. The existence of such thresholds is found to generate both regions of increasing, and regions of decreasing, marginal productivity of increased effort, ability and supporting resources in the presence of stochastic elements in the quality production and assessment processes. The possibility of non-convex regions of the feasible set of assessed quality outcomes is further increased by the endogeneity of several elements of Departmental income, and of its ability to attract and retain able staff, being themselves dependent upon the assessed quality of the Department's output and associated academic reputation.

In the presence of such non-convex regions, the empirical estimates of the extent to which individual academic Departments can improve their performance in each relevant direction depend more critically upon the choice of the analytical technique used to estimate the associated achievement possibility frontier. In particular, the estimates made by DEA, which has been widely used for frontier estimations in the management science and operational research literatures (see Emrouznejad et al, 2008), are found to over-estimate the scope for such improvements by UK Departments of Economics compared to those made by the technique of FDH analysis, which allows for the possibility of both convex and non-convex regions of the achievement possibility set. Nearly twice as many UK Departments of Economics are found to be operating on the efficiency frontier under FDH analysis than under DEA, with the latter's convexity assumption called into question by our above analysis. The existence of non-convex regions of the feasible set in addition tends to increase the possible gains from specialisation in a limited number of directions of achievement, and to increase the importance of issues of price and allocative efficiency compared to those of technical efficiency. Thus, while there may be scope for innovations that shift outwards over time the frontier that is achieved by the most technically efficient Departments, there are also dangers of not adequately recognising the finite nature of the expected achievement possibility set and the trade-offs which need to be made by individual Departments in determining their desired optimal choices over a feasible set of assessed performance that is not guaranteed to be convex.

References

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

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

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

Cameron, A. C. and Trivedi, P.K. (2005). *Microeconometrics: Methods and Applications*, Cambridge University Press, Cambridge.

Charnes, A, Cooper, W.W. and Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2, 429-444.

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

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

Dundar, H. and Lewis, D. R. (1995). Departmental productivity in American universities: economies of scale and scope. *Economics of Education Review*, 14 (2), 119 – 144.

Duprin, 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.

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. J. (1957). The measurement of productive efficiency. *Journal of Royal Statistical Society A*, 120, 253 – 281.

Higher Education Funding Council for England (HEFCE) (2006). *Funding Higher Education in England: How HEFCE Allocates Its Funds*, HEFCE, Bristol.

-------------------- (2008). *RAE2008: Research Assessment Exercise*, http://www.rae.ac.uk.

___________ (2010a). *Guide to funding: How HEFCE allocates its funding*, HEFCE, Bristol.

____________ (2010b). *National Student Survey: Findings and Trends 2006 to 2009*, HEFCE, Bristol.

____________ (2011). *REF2014: Research Excellence Framework*, http:// www. hefce.ac.uk/research/ref.

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, 63-71.

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

Johnes, G. (1999). The management of universities. President's lecture delivered at Annual General Meeting of the Scottish Economic Society. *Scottish Journal of Political Economy*, 46 (5), 505 – 522.

Kumbhakar, S. C. and Lovell, C.A.K. (2000). *Stochastic Frontier Analysis*. Cambridge University Press, Cambridge.

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

Universities Scotland (2008). *Briefing: How is higher education funded?*, Universities Scotland, Edinburgh.