Performance Measurement in UK Universities: Bringing in the Stakeholders' Perspectives Using Data Envelopment Analysis

Cláudia S. Sarrico

Thesis submitted for the Degree of PhD

THE UNIVERSITY OF WARWICK

Supervised by
Prof. Robert G. Dyson

WARWICK BUSINESS SCHOOL

August, 98

TABLE OF CONTENTS

List of Tables

List of Figures

Acknowledgements

I would like to thank various people that helped me throughout my PhD. The most credit goes to Prof. Robert G. Dyson, my supervisor, for his constant support and guidance. I am grateful to Dr. Antreas D. Athanassopoulos, which was my second supervisor during the first part of my PhD studies, for his enthusiasm. I also thank Susanna M. Hogan, our co-author in the paper we wrote together. This thesis has also benefited substantially from the discussions with the rest of the 'Warwick DEA Team': Dr. Emmanuel Thanassoulis, Dr. Rachel Allen, Ali Emrouznejad, Laura E. Read, Ana Santos, Nikolaos Maniadakis, Estelle Shale, and Dr. Victor Podinovski.

I am also grateful to the Rector of Universidade de Aveiro, Prof. Júlio Pedrosa, and Dr. Pedro Vilarinho and Arménio Rego, my lecturers at Universidade de Aveiro, for their initial provocation and for suggesting that I should pursue PhD studies, after my first degree. Their encouragement throughout has been very much appreciated.

As to the people that are dear to me, I would specially like to thank my family and my partner, Julian H. Tice, for their support.

This thesis used data supplied by the Society for Research into Higher Education, by agreement with the Higher Education Statistics Agency. I acknowledge financial support from the Portuguese Foundation for Science and Technology (Fundação para a Ciência e Tecnologia).

Declaration

During my PhD studies the following research papers were written, and conference presentations have taken place.

Research Papers

Sarrico CS and Dyson RG (1998). Performance measurement in UK universities - the institutional perspective. Research Paper, Warwick Business School. Submitted for publication in *J Opl Res Soc.*

Sarrico CS, Hogan SM, Dyson RG and Athanassopoulos AD (1997). Data Envelopment Analysis and University Selection. *Journal of the Operational Research Society* **48**: 1163-1177.

Conference Presentations

Sarrico CS and Dyson RG, *Performance Measurement in UK Universities: The Applicant's Perspective, Young OR9*, Conference of the Operational Research Society, 25-27 March 1996, University of York, UK.

Sarrico CS and Dyson RG, *Medição de Desempenho no Ensino Superior*, 7º Congresso da Associação para o Desenvolvimento da Investigação Operacional, 1-3 April 1996, Universidade de Aveiro, Portugal.

Sarrico CS and Dyson RG, *Performance Measurement in Higher Education*, IFORS 96, 14th Triennal Conference of The International Federation of Operational Research Societies, 8-12 July 1996, Vancouver, Canada.

Sarrico CS and Dyson RG, *DEA and University Selection: A Case Study*, OR38, National Conference of the Operational Research Society, 3-5 September 1996, University of Warwick, UK.

Sarrico CS and Dyson RG, *Performance Measurement in UK Universities: the Institutional Perspective*, OR Young Researchers Forum, 17-18 April 1997, University of Southampton, UK.

Sarrico CS and Dyson RG, *Performance Measurement in UK Universities: the Institutional Perspective*, INFORMS Spring Conference, 4-7 May 1997, San Diego, USA.

Sarrico CS and Dyson RG, *DEA and University Selection: A Case Study*, International Conference on Methods and Applications of Multicriteria Decision Making, 14-16 May 1997, FUCAM, Mons, Belgium.

Sarrico CS and Dyson RG, *Performance Measurement in UK Universities: the Institutional Perspective - a case study*, EURO XV / INFORMS XXXIV Conference, 14-17 July 1997, Barcelona, Spain.

Summary

This thesis is about performance measurement in higher education. It brings in different stakeholders' perspectives on performance measurement in UK universities using data envelopment analysis. The introduction gives the background of the higher education sector in the UK at present and its history. It introduces the drive for performance measurement in higher education, and the motivation for the dissertation. The method data envelopment analysis is then described.

The traditional use of performance indicators and peer assessment is reviewed and the use of DEA, instead of parametric techniques, is justified. The opportunity to use DEA in a somewhat different way than previously is identified.

The novel proposed framework integrates in the same analysis the perspectives of three different levels of stakeholders. Firstly, the perspective of the applicant in the process of choosing a university to apply to; secondly, the perspective of the State that funds and evaluates university performance; and finally the institutional perspective. In the applicant's perspective, the use of DEA in university selection is compared to existing methods. The new approach devised recognises the different values of students and is empirically tested in a case study at a comprehensive school. This chapter clearly deals with a choice problem, and the link with MCDM is first approached. Finally, a comprehensive decision support system that includes DEA for university selection is arrived at.

Then the relationship between the State and higher education over time is described, the current operational model explained and the future trends outlined. In order to measure performance, according to the mission and objectives of the state/ funding councils, a review of their three main remits is undertaken. The contribution of DEA to inform the State/ funding councils in their remit is then discussed. The problem of taking account of subject mix factor in the measurement of performance is dealt with, by linking the input/ output divide by means of virtual weights restrictions.

It is shown how institutions can turn performance measurement to their own benefit, by using it as a formative exercise to understand the different expectations of them, by the two previous external evaluations. A methodology for institutional performance management is proposed that takes into account the external/ internal interfaces: the applicant/ institution, and state/ institution interfaces. The methodology is illustrated with an application to the University of Warwick.

Virtual weights restrictions are widely used in this thesis, a reflection on its uses is offered. The reasons for mainly using virtual weights restrictions instead of absolute weights restrictions are explained. The use of *proportional* weights restrictions is reviewed, and the reasons for using simple virtual weights and virtual assurance regions in this thesis is ascertained. Alternatives to using virtual weights restrictions are considered, namely using absolute weights restrictions with a virtual meaning. The relationship between DEA and MCDM in this domain is elaborated upon.

Several conclusions are arrived at and novel contributions are made to the knowledge of the subject treated: the importance of bringing in the perspectives of different stakeholders in an integrated approach; the contribution of DEA in choice problems; handling subject mix by means of virtual assurance regions; data availability policy is found to be inadequate; a more appropriate way of comparing departments within a university; and the superiority of virtual assurance regions to represent preference structures and link the input-output divide.

Abbreviations

ACC	Academic Cost Centre
API	Age Participation Index
APM	Aggregate Performance Measure
ARI	Assurance Regions of Type I
ARII	Assurance Regions of Type II
ASC	Academic Subject Category
AUCF	Average Unit of Council Funding
BCC	Banker, Charnes and Cooper
BCG	Boston Consulting Group
CCR	Charnes, Cooper and Rhodes
COSHEP	Committee of Scottish Higher Education Principals
CRS	Constant Returns to Scale
CVCP	Committee of Vice-Chancellors and Principals
DEA	Data Envelopment Analysis
DENI	Department for Education in Northern Ireland
DES	Department of Education and Science
DFE	Department for Education
DfEE	Department for Education and Employment
DM	Decision Maker
DMU	Decision Making Unit
FTE	Full Time Equivalent
HEFCE	Higher Education Funding Council for England
HEFCW	Higher Education Funding Council for Wales
HEMS	Higher Education Management Statistics
HEQC	Higher Education Quality Council
HESA	Higher Education Statistics Agency
JPIWG	Joint Performance Indicators Working Group
LEA	Local Education Authority
MCDM	Multicriteria Decision Making
NCIHE	National Committee of Inquiry into Higher Education
NDPB	Non-Departmental Public Bodies
OST	Office for Science and Technology
PCFC	Polytechnics and Colleges Funding Council

PG	Postgraduate
PI	Performance Indicator
QAA	Quality Assurance Agency for Higher Education
RAE	Research Assessment Exercise
SCOP	Standing Conference of Principals
SHEFC	Scottish Higher Education Funding Council
TQA	Teaching Quality Assessment
UAQE	Unit of Assessment of the Quality of Education
UCAS	Universities and Colleges Admission Service
UFC	Universities Funding Council
UG	Undergraduate
UGC	Universities Grants Committee
UKOSA	UK Council for Overseas Student Affairs
UOA	Unit of Assessment
USR	Universities Statistical Record
VRS	Variable Returns to Scale

1. INTRODUCTION

1.1 Introduction

Higher education consists of degree and equivalent courses. It is financed by the governmental budget allocated to the funding councils, following the recommendation of the Secretary of State of the Department for Education and Employment (DfEE). Finance from the Higher Education Funding Councils for England, Scotland and Wales (HEFCE, SHEFC, HEFCW), and the Department for Education in Northern Ireland (DENI) helps meet the costs of teaching, research and related activities in all publicly funded universities and higher education colleges. In addition, institutions undertake paid training, research or consultancy for commercial firms. The research councils provide support for postgraduate studies and research. The funding councils carry out assessments of the education provided by institutions, publish regular reports on their findings and aim to ensure that any serious problems are put right by the university or college concerned. The quality of education in the two Northern Ireland universities is assessed by the HEFCE on behalf of the DENI. The Higher Education Quality Council (HEQC), financed by subscription by institutions, ensured up to March 1997 that satisfactory quality control arrangements were in place. In March 1997 a new agency - The Quality Assurance Agency for Higher Education (QAA) brought together the quality audit and assessment functions together in a single agency.

There are some 90 universities, including the Open University. They are governed by royal charters or by Act of Parliament and enjoy academic freedom. They appoint their own staff, decide which students to admit, provide their own courses, and award their own degrees.

Applications from potential students are usually made through the Universities Colleges Admission Service (UCAS). Overseas students are provided with the services of the British Council and UK Council for Overseas Student Affairs (UKOSA) for this effect.

In this thesis only the universities within the HE sector will be considered, leaving aside a less homogeneous group of institutions that are normally smaller and more specialised.

1.2 UK Universities at Present and their History

The universities of Oxford and Cambridge date from the 12th and 13th centuries, and the Scottish universities of St Andrews, Glasgow, Aberdeen and Edinburgh from the 14th and 15th centuries. All the other universities in Britain were founded in the 19th and 20th centuries (see Table 1-1)[1]. The 1960s saw considerable expansion in the number of new universities, following the publication of the Robbins Report in 1963. The report advocated a substantial increase in the numbers of university places and universities. As a consequence, new universities were established and the Colleges of Advanced Technology upgraded to university status. The development of the public sector polytechnics was initiated in the mid-1960s and there was exceptional growth in student numbers up to 1991. The number of universities also jumped considerably in 1992, when polytechnics and some other higher education establishments were given the freedom to become universities and chose to exercise it.

Table 1-1: Universities in the UK

The Traditional Universities	*The Old Foundations*	Oxford(1264) Cambridge (1284) St. Andrews (1411) Glasgow (1451) Aberdeen (1495) Edinburgh (1583)
	Victorian Expansion: the Federal and Civic Universities	Durham and Newcastle (1832) Belfast (1845) London (1836) Wales (1893) Bristol (1876) Manchester (1880) Dundee (1881) Liverpool (1881) Leeds (1884) Sheffield (1897) Birmingham (1898)
	The 20th Century : London-based Colleges (the second year is the date of attaining independent status from the University of London)	Nottingham (1881, 1948) Reading (1902, 1926) Southampton (1902, 1952) Hull (1927, 1954) Exeter (1922, 1955) Leicester (1918, 1957) Keele (1949, 1962)

The Post Robbins Expansion	*The New Universities : 'Green-Field'*	Sussex (1961) York (1963) East Anglia (1964) Essex (1961) Kent (1964) Lancaster (1964) Warwick (1965) Stirling (1967)
	The New Universities : upgraded Colleges of Advanced Technology	Aston (1966) Bath (1966) Bradford (1966) Brunel (1966) City (1966) Heriot-Watt (1966) Loughborough (1966) Salford (1967) Strathclyde (1964) Surrey (1966)
Unitary System of Higher Education	*The 'Newer' Universities: upgraded Polytechnics and Colleges*	Abertay Dundee (1994) Anglia (1992) Central England, B'ham (1992) Bournemouth (1992) Brighton (1992) West of England, Bristol (1992) Coventry (1992) De Monfort (1992) Derby (1992) East London (1992) Glamorgan (1992) Glasgow Caledonian (1992) Greenwich (1992) Hertfordshire (1992) Huddersfield (1992) Kingston (1992) Central Lancashire (1992) Leeds Metropolitan (1992) Lincolnshire & Humberside (1992) Liverpool John Moores (1992) London Guildhall (1992) Luton (1993) Manchester Metropolitan (1992) Middlesex (1992) Napier (1992) Northumbria (1992) North London (1992) Nottingham Trent (1992) Oxford Brookes (1992) Paisley (1992) Plymouth (1992) Portsmouth (1992) Robert Gordon (1992) Sheffield Hallam (1992) South Bank (1992) Staffordshire (1992) Sunderland (1992) Teeside (1992) Thames Valley (1992) Westminster (1992) Wolverhampton (1992)

When the traditional ancient, civic and federal universities in the UK were founded they were funded from private resources. However, following the formation of the University Grants Committee (UGC) in 1919 as the body responsible for advising the Secretary of State on the funding of universities there was a progressive dominance of public funding and this remained the position for the 'old' universities until 1988.

Initially the polytechnics and colleges established in the mid-1960s received public funding, which was administered by Local Education Authorities in the regions in which they were located.

The position changed radically in 1988 when the Universities Funding Council (UFC) and the Polytechnics and Colleges Funding Council (PCFC) were established under the Education Reform Act 1988. In 1989, between them assumed their formal funding responsibilities for the British universities and the polytechnics and higher education colleges in England. Both councils were Non-Departmental Public Bodies (NDPB) with a high degree of autonomy from the then Department of Education and Science (DES).

In May 1991 a Government White Paper, *Higher Education - a New Framework* proposed a number of substantial changes, the most significant of which was to be the abolition of the so-called 'binary line', between the universities, and the polytechnics and colleges; and the establishment of a unitary system of higher education. In March 1992 the Further and Higher Education Act 1992 was passed by Parliament. The principal organisational change that resulted was the establishment of separate Higher Education Funding Councils for England (HEFCE), Scotland (SHEFC) and Wales (HEFCW). Funding of higher education in Northern Ireland continued to be the responsibility of the Department of Education of the Northern Ireland Office (DENI). The HEFCE, responsible for the funding of most UK institutions, was established formally on 2 June 1992 and it took responsibility for the funding of all higher education institutions in England from 1 April 1993.

The establishment of the new funding agencies brought together the 'old' universities from the UFC sector including the federal universities of London and Wales with its major institutions, the 'new' universities (the former polytechnics) and the higher education colleges in the PCFC sector. The Act also brought into the sector three institutions which were previously funded directly by the Department for Education - the Open University, the Royal College of Art and the Cranfield Institute of Technology. Subsequently, two HE colleges were granted university status and the major institutions of the University of London were given direct access to the HEFCE and treated effectively as if they were independent institutions. Further, two colleges joined the sector from the further education sector. The new sector comprises 115 university institutions in the UK, counting separately the constituent colleges of the federal universities of Wales and London. If Wales and London are counted as single institutions, the total is 90[2].

It should be noted that the Department of Education and Science became the Department for Education (DFE) in May 1992, when the Science Branch of the DES (that part of the department that had been responsible for science policy and the research councils) was transferred to the new Office of Public Service and Science as the Office of Science and Technology (OST). In October 1995, it was created the Department for Education and Employment (DfEE)[3] with the aim of bringing greater coherence in the development of education and employment policies, resulting from the consideration that the nation's future economic competitiveness depends crucially on a workforce well equipped with basic literacy and numeracy skills.

The above developments produced a marked increase in the Age Participation Index (API): particularly in the last few years, as seen in Table 1-2.

Table 1-2: Participation Rates

	Percentage		
	1985/86	1990/91	1995/96
GB API	14	19	30-32

GB API = Under 21 home initial entrants/ average number of 18 and 19 year olds in the population

The funds allocated to higher education, however, did not increase with increased participation, which resulted in a constant erosion of unit funding, as seen by DfEE data in Figure 1-1 (from THESIS data bank[4]).

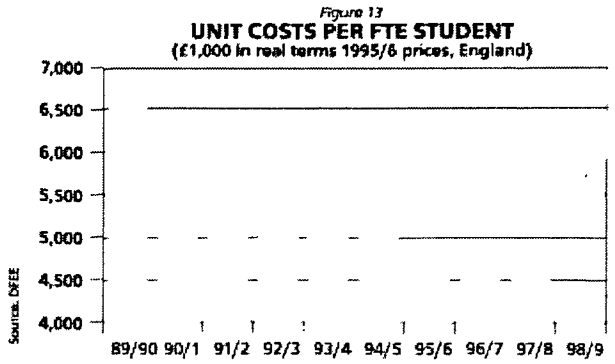

Figure 1-1: Unit Costs per FTE Student

1.3 Motivation for this Thesis

The last years have seen increased pressures for efficiency and effectiveness in the public sector in general, and in the higher education sector in particular. Since the beginning of the eighties funding cuts have been imposed, more accountability demanded and selectivity in the allocation of funds, increasingly dependent on the results of performance assessment exercises, has become common policy.

The review of the literature on performance measurement in universities in chapter 2 shows the potential for the use of data envelopment analysis (DEA), a methodology that seeks to assess.the performance of homogeneous organisational units.

This thesis is concerned with establishing a theoretical framework for the measurement of performance in universities using DEA, considering not only the evaluative side of the exercise, but also as a tool for the enhancement of performance and organisational change. For this purpose, our approach explicitly takes into consideration the existence of different universities' stakeholders and their purposes for measuring performance, against a background of what can currently be said about their mission and objectives. We also discuss some issues that emerge from the interaction of the different stakeholders perspectives.

The adequacy of current performance measures and the possible inappropriate behavioural responses they may generate is discussed, and so is the data availability policy currently in use.

Finally, a discussion of the use of weights restrictions in DEA, that emerge from its use in the applications of the framework is presented.

1.4 The Method: Data Envelopment Analysis

In 1957 a paper was published in the journal of the Royal Statistical Society by M. J. Farrell on *The Measurement of Productive Efficiency*[5]. This paper provided the background for data envelopment analysis.

DEA emerged during a study by E. Rhodes, under the supervision of A. Charnes and W. W. Cooper, to evaluate the performance of an educational programme (PFT - Program Follow Through) for disadvantaged pupils in the USA. The analysis involved comparing the performance of a set of school districts that were participating in PFT and a set that was not. It was the challenge of estimating the relative efficiency of the schools involving multiple outputs and inputs, without information on prices, that resulted in the formulation of the ratio form of DEA, known by the initials of its developers - CCR (Charnes, Cooper, and Rhodes). It was first published in 1978 in the European Journal of Operational Research[6]. Thus DEA began as a tool to measure efficiency of public sector organisations.

1.4.1 The Concept

DEA is a method used to estimate the efficiency of homogeneous organisational units, called decision making units (DMUs), that use the same inputs to produce the same outputs. DEA takes the observed input and output values to form a production possibility space, against which the individual units are compared to determine their efficiencies. The output efficiency of a unit measures the amount by which the output of that unit can be increased without the need to increase the inputs. The input efficiency is defined likewise.

The CCR model measures the efficiency of target unit j_o relative to a set of peer units:

CCR Model

$$e_0 = \max_{u,v} \frac{\sum_{r=1}^{s} u_r y_{rj_0}}{\sum_{i=1}^{m} v_i x_{ij_0}}$$

$$s.t. \quad \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leq 1, \quad j = 1,\ldots,n$$

$$u_r, v_i \quad \geq \varepsilon, \ \forall r \, and \, i$$

where

y_{rj}	=	amount of output r from unit j,
x_{ij}	=	amount of input i from unit j,
u_r	=	the weight given to output r,
v_i	=	the weight given to input i,
n	=	the number of units,
s	=	the number of outputs,
m	=	the number of inputs,
ε	=	a positive non-Archimedean infinitesimal.

It is assumed that there are n DMUs to be evaluated. Each DMU consumes varying amounts of m different inputs to produce s different outputs. The CCR model translates into the following: unit j_o is said to be efficient (e_o=1) if no other unit or combination of units can produce more than unit j_o on at least one output without producing less in some other output or requiring more of at least one input.

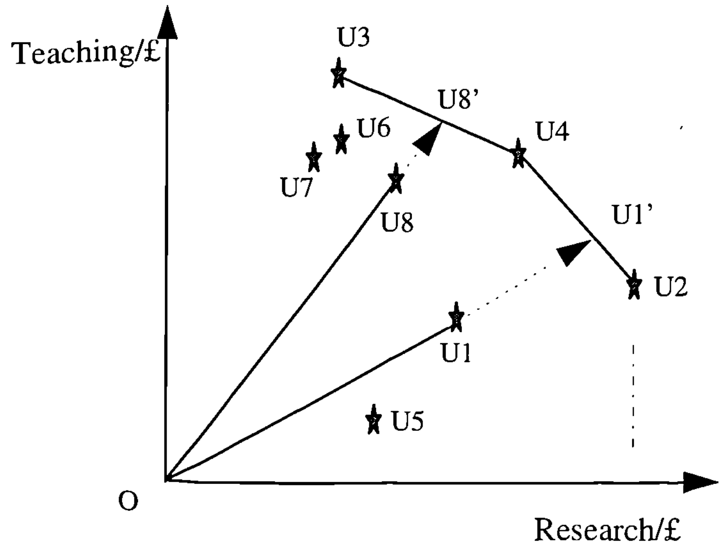

Figure 1-2: DEA - a graphical representation

A graphical illustration of the concept is presented in Figure 1-2. Assuming a simple model of a university production process, there are two outputs, teaching and research, normalised by unit of funding. The solid line in Figure 1-2 represents the frontier derived by DEA from data on a population of universities, each producing different amounts of teaching and research outputs per unit of input. *U3*, *U4* and *U2* are efficient units, which are the comparators for the inefficient units they *envelop*. DEA optimises on each individual observation with an objective of calculating a discrete piecewise frontier determined by the set of efficient units (*U3*, *U4*, and *U2*). It contrasts with parametric analysis, such as regression, in which the single optimised regression equation is assumed to apply to each DMU. DEA, unlike parametric methods, does not require any assumption about a functional form that relates independent to dependent variables. It calculates a maximal performance measure for each DMU relative to all other DMUs in the observed population, with the only requirement that each DMU will lie on or below the frontier (i.e. $e \leq 1$). Each inefficient DMU, not on the frontier is scaled against a linear combination of the DMUs on the frontier, e.g. *U8'*, and its measure of efficiency is given by the ratio $e = \overline{OU8}/\overline{OU8'}$.

DEA produces only *relative* efficiency measures, since the calculations are generated from actually observed data for each DMU. The relative efficiency of each DMU is calculated *in relation* to all the other DMUs, using the observed values for the outputs and inputs of each DMU. DEA maximises the relative efficiency score of each DMU, subject to the condition that the set of weights obtained in this manner for each DMU must also be feasible for all the other DMUs included in the calculation. The produced piecewise empirical production frontier represents the best-practice production frontier - the maximum output empirically obtainable from any DMU in the observed population, given its level of inputs.

For each inefficient DMU, one that is enveloped by the efficient set, DEA identifies the sources and level of inefficiency for each of the inputs and outputs. The level of inefficiency is determined by comparison to a single or linear combination of other efficient units that use the same level of inputs and produce the same or a higher level of outputs. This is achieved by increasing some outputs (or decrease some inputs) without worsening the other inputs or outputs. The potential improvement for each DMU does not necessarily correspond to observed performance, but is indicative of potential improvements based on the best-practice performance of 'comparable' DMUs located in the efficient frontier.

The CCR model assumes constant returns to scale (CRS). In 1984 Banker, Charnes and Cooper[7] proposed the BCC model, which relaxed the original CRS requirement of the CCR ratio model. The BCC model makes it possible to investigate local returns to scale, under variable returns to scale (VRS) assumption. Under VRS assumption, DEA can complement information about average returns to scale with DMU-specific scale efficiency for each DMU on the frontier. A detailed formulation of these models is provided in the next section.

1.4.2 Basic DEA Models

The CCR model is a linear fractional model and it needs to be transformed to an ordinary linear programme to be solved. This can be done by scaling either the denominator or the numerator of the objective function equal to a constant such as 1. The equivalent linear programming models (*multipliers formulation*), and their duals (*envelopment formulation*), are as follows:

Table 1-3: Input Oriented CRS

Primal		*Dual*	
$e_0 = \max \sum_{r=1}^{s} u_r y_{rj_o}$		$e_0 = \min h_0 - \varepsilon \sum_{i=1}^{m} s_i^- - \varepsilon \sum_{r=1}^{t} s_r^+$	
$s.t.$		$s.t.$	
$\sum_{i=1}^{m} v_i x_{ij_o}$	$= 1$	$h_0 x_{ij_o} - \sum_{j=1}^{n} \lambda_j x_{ij} - s_i^-$	$= 0, \forall i$
$-\sum_{i=1}^{m} v_i x_{ij} + \sum_{r=1}^{s} u_r y_{rj}$	$\leq 0, \forall j$	$\sum_{j=1}^{n} \lambda_j y_{rj} \quad - s_r^+$	$= y_{rj_0}, \forall r$
$-v_i$	$\leq -\varepsilon, \forall i$	$\lambda_j, \quad s_i^-, \quad s_r^+$	$\geq 0, \forall j, r, i$
$-u_r$	$\leq -\varepsilon, \forall r$		

Table 1-4: Output Oriented CRS

Primal		*Dual*	
$\frac{1}{e_0} = \min \sum_{i=1}^{m} v_i x_{ij_o}$		$\frac{1}{e_0} = \max \theta_0 + \varepsilon \sum_{i=1}^{m} s_i^- + \varepsilon \sum_{r=1}^{t} s_r^+$	
$s.t.$		$s.t.$	
$\sum_{r=1}^{s} u_r y_{rj_o}$	$=1$	$\sum_{j=1}^{n} \lambda_j x_{ij} + s_i^-$	$= x_{ij_0}, \forall i$
$\sum_{i=1}^{m} v_i x_{ij} - \sum_{r=1}^{s} u_r y_{rj}$	$\geq 0, \forall j$		
v_i	$\geq \varepsilon, \forall i$	$\theta_0 y_{rj_0} - \sum_{j=1}^{n} \lambda_j y_{rj} + s_r^+$	$= 0, \forall r$
u_r	$\geq \varepsilon, \forall r$	$\lambda_j, \quad s_i^-, \quad s_r^+$	$\geq 0, \forall j, i, r$

From the *envelopment formulation*, we can see that in the *input oriented* models, one focuses on maximal movement towards the frontier through proportional reduction of inputs (by a factor h_0), whereas in the *output oriented* models, one focuses on maximal movement via proportional augmentation of outputs (by a factor θ_0). The choice of model will depend on which factors under consideration are more easily controlled by the DMU.

The CRS models yield an evaluation of overall efficiency, and identify the sources and estimate the amount of inefficiency. The VRS models, on the other hand, distinguish between technical and scale inefficiencies by estimating pure technical efficiency at the given scale of operation, and identifying whether increasing, decreasing, or constant returns to scale possibilities are present. Under VRS, the linear models are as below:

Table 1-5: Input Oriented VRS

Primal		Dual	
$e_0 = \max \sum_{r=1}^{s} u_r y_{rj_o} + \omega_0$		$e_0 = \min h_0 - \varepsilon \sum_{i=1}^{m} s_i^- - \varepsilon \sum_{r=1}^{t} s_r^+$	
$s.t.$		$s.t.$	
$\sum_{i=1}^{m} v_i x_{ij_o}$	$= 1$	$h_0 x_{ij_o} - \sum_{j=1}^{n} \lambda_j x_{ij} - s_i^-$	$= 0, \forall i$
$-\sum_{i=1}^{m} v_i x_{ij} + \sum_{r=1}^{s} u_r y_{rj} + \omega_0$	$\leq 0, \forall j$	$\sum_{j=1}^{n} \lambda_j y_{rj} \quad - s_r^+$	$= y_{rj_0}, \forall r$
$-v_i$	$\leq -\varepsilon, \forall i$	$\sum_{j=1}^{n} \lambda_j$	$= 1$
$-u_r$	$\leq -\varepsilon, \forall r$		
ω_0	*free*	$\lambda_j, \quad s_i^-, \quad s_r^+$	$\geq 0, \forall j, r, i$

Table 1-6: Output Oriented VRS

Primal		Dual	
$\frac{1}{e_0} = \min \sum_{i=1}^{m} v_i x_{ij_o} - \omega_0$		$\frac{1}{e_0} = \max \theta_0 + \varepsilon \sum_{i=1}^{m} s_i^- + \varepsilon \sum_{r=1}^{t} s_r^+$	
$s.t.$		$s.t.$	
$\sum_{r=1}^{s} u_r y_{rj_o}$	$= 1$	$\sum_{j=1}^{n} \lambda_j x_{ij} + s_i^-$	$= x_{ij_0}, \forall i$
$\sum_{i=1}^{m} v_i x_{ij} - \sum_{r=1}^{s} u_r y_{rj} - \omega_0$	$\geq 0, \forall j$	$\theta_0 y_{rj_0} - \sum_{j=1}^{n} \lambda_j y_{rj} \quad + s_r^+$	$= 0, \forall r$
v_i	$\geq \varepsilon, \forall i$		
u_r	$\geq \varepsilon, \forall r$	$\sum_{j=1}^{n} \lambda_j$	$= 1$
ω_0	*free*	$\lambda_j, \quad s_i^-, \quad s_r^+$	$\geq 0, \forall j, i, r$

To illustrate what happens when the different models are used, seven example DMUs (*P1,..., P7*) are shown in Figure 1-3. An inefficient DMU can be made fully efficient by projection onto a point $(\hat{X}_0, \hat{Y}_0)$ on the envelopment surface. The particular point of projection selected will be dependent not only on the returns to scale assumption, but also on the orientation of the model, as shown for *P5*.

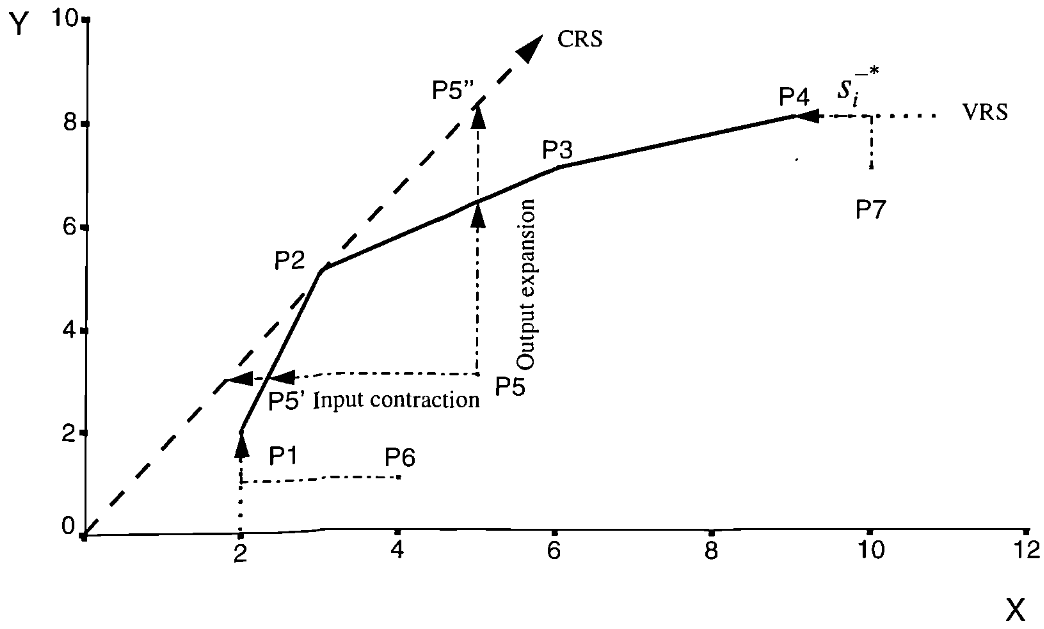

Figure 1-3: Example DMUs

The difference between the CRS and VRS models is the convexity constraint $\left(\sum_{j=1}^{n}\lambda_j = 1\right)$ in the envelopment formulation, and the variable ω_0 in the multiplier formulation. The presence of the convexity constraint in the VRS models diminishes the feasible region for the DMUs. The result is an increase in the number of efficient DMUs: under CRS only DMU 2 is efficient, whereas under VRS DMUs 1 to 4 are. As for the variable ω_0, it will define the y-intercept corresponding to the segments of the piecewise frontier, and thus define the local returns to scale applying to each target unit, assuming VRS. See Table 1-7 for unique optimal solutions. Multiple optimal solutions are addressed in Banker and Thrall (1992)[8]. Under CRS, $\omega_0 \equiv 0$, and thus the frontier passes through the origin.

Table 1-7: Local returns to scale under VRS

Increasing	*Constant*	*Decreasing*
$\omega_0^* > 0$	$\omega_0^* = 0$	$\omega_0^* < 0$

In an *input oriented model*, the variable h_0 represents the reduction applied to all inputs of DMU j_0 to improve efficiency. The presence of the non-Archimedean ε in the *multipliers formulation* objective function allows for the minimisation of h_0 to pre-empt the optimisation of the slacks. The correct algorithmic implementation requires a two-stage approach, which is the case of the software[9] used in this thesis. Thus, the optimisation is computed in a two-stage process, with maximal reduction of inputs being achieved first; and then, in the second stage, movement onto the frontier is achieved via the slack variables ($s_i^- \, and \, s_r^+$), as for *P6*. That is: a DMU is efficient if and only if the two conditions are satisfied h_0=1, and all slacks are zero. The non-zero slacks and the value of $h_0 \leq 1$ identify the sources of inefficiencies. Under VRS, DMUs 1 to 4 are efficient and define the piecewise linear envelopment surface, the input shrinkage and all slacks being zero projects DMU 5 onto *P5'* lying on the segment *P1P2*.

In an *output oriented model* the objective is to achieve via the variable θ_0 as much expansion of the outputs as possible. Additionally, as before, that might not be enough to achieve efficiency. Thus for DMU 7, the efficient projection onto the VRS frontier requires both an output augmentation by θ_0^* *and* the slack value s_i^{-*}, as seen in Figure 1-3.

Under CRS the previous discussion for the role of the non-Archimedean constant still applies, i.e. the proportional input reduction (output expansion) may not be enough to achieve efficiency. This is not obvious for a two-dimensional example as in Figure 1-3, but in a multiple-input multiple-output example, slacks may be present and necessary to reach the efficient frontier. It may even be the case that simultaneous augmentation of outputs *and* reduction in inputs occurs, as seen by the expressions for the obtained targets below.

For input oriented models:

$$\hat{x}_{ij_0} = \sum_{j=1}^{n} \lambda_j^* x_{ij} = h_0^* x_{ij_o} - s_i^{-*}, \forall i$$

$$\hat{y}_{rj_0} = \sum_{j=1}^{n} \lambda_j^* y_{rj} = y_{rj_0} + s_r^{+*}, \forall r$$

For output oriented models:

$$\hat{x}_{ij_0} = \sum_{j=1}^{n} \lambda_j^* x_{ij} = x_{ij_0} - s_i^{-*}, \forall i$$

$$\hat{y}_{rj_0} = \sum_{j=1}^{n} \lambda_j^* y_{rj} = \theta_0^* y_{rj_0} + s_r^{+*}, \forall r$$

In summary, either under CRS or VRS a unit is considered efficient in an input oriented model if and only if it is efficient in the corresponding output oriented model. However, although the two orientations provide the same envelopment surfaces, an inefficient DMU is projected to different points on the frontier under the input and output orientations, and will possibly have different efficient units as peers. On the other hand, the relationship between the CRS and VRS models is different. If a DMU is characterised as efficient in the CRS model, it will also be characterised as efficient in the VRS model, the converse is not necessarily true.

In conclusion, the choice of a DEA formulation requires a choice of orientation - minimisation of inputs or maximisation of outputs, and a scale assumption - constant returns to scale or variable returns to scale. Other models, namely additive and multiplicative, can be found in the DEA reference text by Charnes et al, (1994)[10].

1.4.3 Incorporating Judgement or A Priori Knowledge

One of the recognised advantages of the DEA methodology is that a priori specification of the weights (multipliers formulation) is not required, and each DMU is evaluated in the best possible light. However, the absolute freedom given to the units to choose their own sets of weights can have undesirable consequences. This freedom can lead to some units appearing efficient through a judicious choice of weights, rather than good performance. One approach to this problem is to impose weights restrictions to curb the complete freedom under classical DEA, incorporating expert opinion, the decision-maker's preferences, or other judgement into the analysis. Restrictions can be directly applied to the multipliers, normally named *absolute* weights restrictions, or applied to the virtuals, named *virtual* weights restrictions. The virtual inputs and outputs of a DMU reveal the relative contribution of each input and output to its efficiency rating. Virtual input/ output is the product of the input/ output level and optimal weight for that input/ output. The higher the level of virtual input/ output, the more important that input/ output is in the efficiency rating of the DMU concerned. Therefore use of virtual inputs and outputs can help to identify strong and weak areas of performance.

Proposed techniques include to impose bounds on ratios of multipliers (Thompson et al, 1986[11]), imposing upper and lower bounds on individual multipliers (Dyson and Thanassoulis[12], 1988), appending proportional virtual weights restrictions (Wong and Beasley[13], 1990); and transforming input-output data to simulate weights restrictions (Charnes et al.,1990[14]). For more on weights restrictions see Allen et al. 1997[15].

In this thesis, *virtual* weights constraints applied to the target unit have been used to translate a priori judgements, although the most appropriate way to implement virtual weights restrictions is not fully resolved. In chapter 6 a discussion of the application of virtual weights restrictions is undertaken.

1.5 Outline of this Thesis

In this chapter, we have provided an overview of the thesis: the background of the UK university sector and the context they operate in, the motivation for writing the thesis, and a description of the methodology adopted.

In Chapter 2, *Performance Measurement in Universities*, we first undertake a review of the use of performance measurement in UK universities and the methods used thus far. Secondly, we justify the use of the DEA methodology instead of other methods, and critically review the published research on the use of DEA for performance measurement in universities. Finally we propose a new framework for the use of DEA in performance measurement in UK universities that takes into account who is doing the assessment, for what purpose, and which criteria are used. Bringing into the assessment three different levels of stakeholders interested in the performance of universities: the State, the institution, and the applicant, gives a global picture of UK universities performance measurement, and how the different perspectives interact.

Chapter 3, *The Applicant's Perspective*, is concerned with the process of university selection by applicants to UK universities. The appropriateness of popular league tables, such as the *Times*', in guiding applicant's choices is considered, and a case study of applicants at a comprehensive school reported on. The contribution of DEA as a decision support technique, which can produce customised individual league tables to inform the potential student in his/ her choice, is illustrated. Finally it explores the design of a decision support system for university selection using DEA with informed judgement.

Chapter 4, *The State Perspective*, reviews the relationship between the State and the higher education sector, and how the funding councils implement state policy. The contribution of DEA to inform the funding councils on their remit of assessing quality and allocating resources is ascertained. Some DEA models for measuring the performance of the university sector are proposed, and an application to the academic year 1995/96 is described.

In Chapter 5, *The Institutional Perspective*, presents the institutions 'sandwiched' between the external assessment by the two former classes and their own objectives. The contribution of DEA to inform management is explored. First, a theoretical framework for institutional performance management is described, that takes into account the interfaces between its internal and external environments. Then an application of the methodology to the University of Warwick is described.

Chapter 6, *The Use of Virtual Weights Restrictions in DEA*, discusses the use of virtual weights restrictions in DEA. We have as the point of departure the work of Wong and Beasley[13], which first proposed restricting the virtuals, instead of the absolute multipliers. The implications of their proposed three approaches for imposing *proportional* virtual weights restrictions are considered. Then it discusses the use of virtual weights restrictions in this thesis, which is somewhat different from the previous approaches. The chapter concludes with some general principles for the implementation of virtual weights restrictions.

The final chapter summarises the findings of this thesis, the contribution it makes for the analysis of performance measurement in HE, the limitations of the material, and also points to further areas of research.

1.6 References

[1] Higher Education Funding Council for England (1994). Overview of Recent Developments in HE, M2/94. HEFCE: Bristol.

[2] Committee of Vice-Chancellors and Principals of the Universities of the United Kingdom (1997). *Higher Education Statistics, Autumn 1997.* CVCP: London.

[3] Great Britain, Department for Education and Employment (1995). *New Department Will Sharpen Focus On Basic Skills, 237/95.* DfEE: London.

[4] Times Higher Education Supplement Information Service (1997). *Higher Education Trends: Finance*. Times: London.

[5] Farrell MJ (1957). The measurement of productive efficiency. *Jl R Statist Soc A* **120**: 253-281.

[6] Charnes A, Cooper WW and Rhodes E (1978). Measuring the efficiency of decision making units. *Eur J Opl Res* **2**: 429-444.

[7] Banker RD, Charnes A, Cooper WW (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Mgmt Sci* **30**: 9 1078-1092.

[8] Banker RD and Thrall RM (1992). Estimation of returns to scale using data envelopment analysis. *Euro J Opl Res* **62**: 74-84.

[9] (1996). *Warwick Windows DEA, Version 1.02.* Warwick Business School: Coventry.

[10] Charnes A, Cooper WW, Lewin AY and Seiford LM (Eds) (1994). *Data Envelopment Analysis*. Kluwer: Boston.

[11] Thompson RG, Singleton FD, Thrall RM and Smith BA (1986). Comparative site evaluations for locating a high-energy physics lab in Texas. *Interfaces* **16**: 35-49.

[12] Dyson RG and Thanassoulis E (1988). Reducing weight flexibility in data envelopment analysis. *J Opl Res Soc* **39**: 563-576.

[13] Wong Y-HB and Beasley JE (1990). Restricting weight flexibility in data envelopment analysis. *J Opl Res Soc* **41**: 829-835.

[14] Charnes A, Cooper WW, Huang ZM and Sun DB (1990). Polyhedral cone-ratio DEA models with an illustrative application to large commercial banks. *J Econometrics* **46**: 73-91.

[15] Allen R, Athanassopoulos AD, Dyson RG and Thanassoulis E (1997). Weights restrictions and value judgements in data envelopment analysis: evolution, development and future directions. *Ann Ops Res* **73**: 13-34.

2. PERFORMANCE MEASUREMENT IN UNIVERSITIES

2.1 Introduction

The last years have seen increased pressures for efficiency and effectiveness in the public sector in general, and in the higher education sector in particular. Since the beginning of the eighties funding cuts have been imposed, more accountability demanded and selectivity in the allocation of funds, increasingly dependent on the results of performance assessment exercises, has become common policy.

Cave et al (1988)[1] report pressures for accountability of the UK higher education system in the same way as other public services that receive money from the Government. In 1985 the Jarratt Report[2], issued by the Committee of Vice-Chancellors and Principals (CVCP) in conjunction with the Government and the Universities Grant Committee (UGC), suggested that the universities should work to clear objectives; achieve value for money; and have strong management and planning structures. The Government would thus be concerned to ensure that universities, as much as other public sector institutions, met objectives determined outside themselves, and demonstrated they achieved these goals. In 1985, the Green Paper *The Development of Higher Education into the 1990s*[3] advocated the development of performance indicators (PIs). By 1987 the White Paper, *Higher Education: Meeting the Challenge*[4] proposed the replacement of the UGC by the Universities Funding Council (UFC), that became responsible for the distribution of funds among universities under new contract arrangements, and a new Polytechnics and Colleges Funding Council (PCFC) was established, which would also submit public sector institutions to a system of contract funding. Performance would be monitored in accordance with those contracts. The White Paper also made recommendations for quality and efficiency, including more selectively funded research, 'targeted with attention to prospects for commercial exploitation', and it favoured the development and use of PIs. In 1987/88 the provision of Government funds already depended crucially on evidence of real progress in implementing and building upon the proposed changes.

The 1991 White Paper, *Higher Education: A New Framework*[5], not only proposed the end of the binary line, the establishment of the new funding councils, research funding to be allocated entirely on a selective basis; but also introduced a new emphasis on quality. It suggested that quality audit should be the responsibility of the institutions and quality assessment that of the funding councils. The latter would inform funding and be based on two approaches: first, the use of PIs and calculations of value added; and, second, the external judgements on the basis of direct observations.

Currently there is selectivity in the distribution of resources and rationalisation, and, when appropriate, even the closure of small departments as an incentive to better financial management and improved standards of teaching.

2.1.1 The Introduction of Performance Indicators

The Jarratt Committee suggested a major shift in evaluation methods from subjective peer review to performance indicators as a means of performance measurement. It recommended explicit quantitative as well as qualitative judgements.

In 1986 the Government, CVCP and UGC agreed on a Concordat[6] that required changes to be met for further finance to be released to universities, and outlined a set of possible PIs. Then in 1987 their Joint Working Party listed 39 PIs for publication[7] relating to both inputs and outputs in teaching and research. They were published by the CVCP under the name *University Management Statistics and PIs*[8], for British universities. The ninth and final annual volume, still for the old universities only, was published with data (by now it listed 54 measures) for the 1992/93 academic year[9]. In 1992 the Joint Performance Indicators Working Group (JPIWG) of the Committee of Vice-Chancellors and Principals of the Universities of the UK (CVCP); the Standing Conference of Principals (SCOP); and the Committee of Scottish Higher Education

Principals (COSHEP); was established, to build on earlier work and develop performance measures and indicators of both teaching and research, for the new merged university sector. By 1993 it circulated the national totals for the number of publications for each cost centre in 20 categories and averages for three indicators[10]. However, institutional figures were not published and the exercise was never repeated.

In 1995 another group - Higher Education Management Statistics (HEMS)[11] - was created by representative bodies of the institutions (CVCP, SCOP and COSHEP), which was invited by the funding councils to take forward the work of the JPIWG. Later, in a consultative document[12], it alerted institutions to the macro and institutional statistics - note that they are no longer called PIs - that were planned to be published by the Higher Education Statistics Agency (HESA) in 1996. It invited suggestions on health warnings and caveats. By then it was no longer intended there should be a successor to the *Universities Management Statistics and Performance Indicators* published for the old university sector by the University Statistical Record (USR). Perhaps not surprisingly, only the macro statistics[13], since the academic year 1994/95, have been published to this date. Publication of the institutional statistics for the academic year 1995/96 has been announced by HESA, but has not been completed yet.

This delay on agreed statistics to be published by HESA was noted in a letter by the DfEE to HEFCE:

> *'would like to see progress with the setting of performance indicators for the higher education sector as a whole' (DfEE, 1996)*[14]

The remit of the state to inform the 'clients' of higher education and the consequences for data availability policy will be discussed in Chapter 4.

This difficulty of finding appropriate measures to portray institutional performance stems from the rise of the quality movement (for a review see Cave et al, 1997)[15], and the realisation that some qualitative issues were not being dealt with by the, often crude, quantitative measures that had been used in the past. This argument was especially true in the more heterogeneous unitary system of higher education, with institutions with very different characteristics and missions. More holistic approaches were favoured, such as the peer review of teaching and research, in the form of the teaching quality and research assessment exercises, described in the next section.

2.1.2 Peer Assessment: Teaching and Research Assessment Exercises

In 1985 in a letter[16] to universities explaining the new resource allocation procedure, the UGC referred to research PIs but lamented that the availability of few indicators of teaching performance would not permit a systematic external assessment of teaching quality to be made. It suggested that a suitable methodology for taking account of teaching quality could be used in due course.

In May 1986 the UGC research ratings by 'cost centres' were sent to universities[17] following a four-point scale: star, for outstanding, the highest rating; A+, above average; A, average, and A-, below average. In 1989 the research selectivity exercise (as it was then called) classified departments into one of five (instead of four) categories. By then the published ratings[18] were already being used to inform calculations of grant to universities. In 1992 a third research assessment exercise - RAE[19] (now with yet another different name) took place to allocate research funds for the new merged university sector. Subsequent funding[20] translated the one to five rating into a funding scale of zero to four, bringing increased selectivity into the funding system. For the 1996 RAE[21], the new emphasis on quality rather than quantity is translated in a key change: the details of up to four publications per researcher would be requested for scrutiny, but not the total number of publications. Selectivity is further increased by translating the seven point scale (1, 2, 3b, 3a, 4, 5 and 5*) into a funding scale of zero for the lower two ratings; each subsequent rating point attracting a weight 50% greater than the previous point, with an 20% premium for the top rating[22], instead of the previously used linear scale.

The assessment of the quality of teaching, now known as the teaching quality assessment exercise, took more time to develop. The Scottish council's approach to quality assessment in 1992 included a framework that covered 11 dimensions, and all institutions would be visited[23]. In 1993 the English council outlined the purposes and methods of its quality assessment[24]. The purposes included: to ensure that higher education was of satisfactory quality or better; to encourage improvements in quality; and to inform funding and reward excellence. Subject areas would be assessed within three categories: excellent, satisfactory and unsatisfactory. Assessment would comprise scrutiny of an institution's self-assessment supported by statistical indicators; together with an assessment visit, in the case of institutions claiming excellence, all whose self-assessment suggested they might be unsatisfactory, and a sample of others. In 1994, the Barnett Report[25] proposed some changes to the HEFCE method of assessment (used also by HEFCW, and by HEFCE on behalf of DENI), including universal visiting, summative judgement at the threshold level, and the framing of recommendations for improvement within a limited number of dimensions, so as to produce a profile. Thus HEFCE announced in 1994 a new method of assessment[26], which is currently in use. It includes the proposed universal visiting, a framework based on six core aspects of provision, profiles structured within four grades, and an overall summative judgement at the threshold level.

On 27 March 1997 a new agency - The Quality Assurance Agency for Higher Education (QAA) - that brought together the quality audit and assessment functions was established. It took over the work of the participating funding councils' quality assessment divisions and all the functions of the Higher Education Quality Council (HEQC). The institutions are also represented by their representative bodies: Committee of Scottish Higher Education Principals (COSHEP); the Committee of Vice-Chancellors and Principals of the Universities of the UK (CVCP); the Heads of Higher Education Institutions in Wales (HHEW); the Standing Conference of Principals (SCOP); and the participating funding councils: the Department for Education Northern Ireland (DENI); the Higher Education Funding Councils for England and Wales (HEFCE and HEFCW).

The QAA has recognised that opinions are still divided over the value of audit and assessment. However, it warned that, like it or not, institutions would be judged and funded by their measured performance:

> *'As a sector, we must take steps to have the largest say in what is to be measured and how.' (HEFCE, 1997)*[27]

In conclusion, the measurement of performance is here to stay, but institutions and their representatives want to have a say in how it is undertaken, instead of having an external methodology imposed on them.

In the next section the reasons for using DEA as a methodology of performance assessment in universities are discussed.

2.1.3 Why Use Data Envelopment Analysis

As far back as 1987, Smith and Mayston[28] argued that the pursuit of efficiency in the public sector, by publishing performance indicators for individual agencies, was fraught with the problem of interpreting them in isolation. They then suggest the use of DEA to show how the data underlying performance indicators can be used to generate a single measure of efficiency for an agency, which additionally adjusts for differences in the environment that different agencies face. Cave et al (1988)[1] in their review of the use of performance indicators lament early on in the book the little use of DEA, a promising tool, which had not been fully explored in higher education, despite its theoretical attractions. They suggest that as more data becomes available, studies showing consistent results when the inputs and outputs identified and other assumptions made are changed, that DEA could have a greater impact upon assessment in higher education. On similar lines, Smith (1990)[29] assesses developments of the use of performance indicators in the public sector in relation to work in the private sector. It points to regression analysis and DEA as analytic techniques available to interpret performance indicators.

The use of regression analysis to evaluate performance in UK higher education has been used particularly by a group of researchers at Lancaster University - G. Johnes, J. Johnes and J. Taylor. Work has been done on the research output of universities (Johnes, 1988[30]; Johnes, 1990[31]; Johnes et al, 1993[32]; Taylor, 1994[33]; Taylor, 1995[34]), progression and completion rates (Johnes and Taylor, 1989[35]; Johnes, 1990[36]), student attainment (Johnes and Taylor, 1987[37]), employment and further study (Johnes et al, 1987[38]; Johnes and Taylor, 1989[39]), and unit costs (Johnes, 1990[40]). Most of their findings are collected in Johnes and Taylor (1990)[41], where they also conclude that it would be interesting and potentially valuable to compare the results of the regression approach with those obtained using DEA. Additionally, Johnes's review (1992)[42] of the use of performance indicators in higher education, recognises the advantage of DEA as a tool that does not involve imposing arbitrary weights. 'The method holds much promise.' DEA has recently been used to explore the research performance of economic departments of UK universities (Johnes and Johnes, 1993[43]; Johnes and Johnes, 1995[44]; Johnes, 1995[45]) taking advantage of the data available following the research assessment exercises. However, unlike what Cave et al (1988)[1] had predicted, data on teaching performance did not become as readily available.

DEA has been compared to ratio analysis (Thanassoulis et al, 1996[46]), regression analysis (Thanassoulis, 1993[47]), and subjective peer review (Doyle et al, 1996[48]). DEA, unlike performance indicators, which are difficult to interpret in isolation, generates an aggregate performance measurement (APM) for the unit under analysis. DEA can also set targets so that units become efficient. Unlike regression analysis, DEA allows for multiple inputs and outputs. It is a non-parametric method, not requiring the user to hypothesise a mathematical form for the production function. DEA measures performance against a frontier of best observed practice, rather than average performance. DEA can also be used as an idealised model, where units judge themselves, consistent with policy constraints set outside themselves, as in the case of peer review. In fact, peer review, is normally quite an expensive method of performance assessment, which might profit from the support of quantitative methods like DEA, specially in models that incorporate measures of quality.

In the next section the use of DEA in the context of university performance is reviewed.

2.2 Using DEA to Measure Performance in Universities

DEA has been used to compare the relative efficiency of public and private universities in the USA (Ahn, 1987[49]; Ahn et al, 1988[50]; Ahn et al, 1989[51]; Ahn and Seiford, 1993[52]); efficiency and perceived quality of the top national universities in the USA (Breu and Raab, 1994[53]), and the cost and value-added efficiencies of universities in the UK (Athanassopoulos and Shale, 1997[54]). DEA has also been used to compare the overall efficiency of university departments of accounting (Tomkins and Green, 1988[55]), chemistry and physics (Beasley, 1990[56]; Beasley, 1995[57]); the research performance of departments of economics (Johnes and Johnes, 1993[43]; Johnes and

Johnes, 1995[44]; Johnes, 1995[45]), and business schools (Doyle et al, 1996[48]) in UK universities; and business and economics departments of Danish universities (Olesen and Petersen, 1995[58]); and university departments of the same university (Sinuany-Stern et al, 1994[59]) in Israel. Wilkinson (1991)[60] uses a modified version of DEA to deal with incomplete data sets, either because the data is not available, or because of 'specialisation' of a decision making unit (DMU). Doyle and Green (1994)[61] even suggest its use in student and staff appraisal in higher education. Table 2-1 summarises the DEA models used in these studies.

Table 2-1: The Use of DEA in University PM

Study	Purpose	Inputs	Outputs	Model
		Comparing Universities		
Ahn (1987)[49], Ahn et al (1988)[50]	Compare the efficiency of 81 private and 80 state doctoral-granting universities, with and without medical schools in the USA for the academic year 1985-86.	Instructional expenditures Overhead expenditures Physical investments	Undergraduate FTEs Graduate FTEs Federal Research Grants and Contracts	CCR and BCC
Ahn et al (1989)[51]	Compare the efficiency of 37 state senior colleges and universities in Texas for five academic years 1981-1985.	Total Faculty Salaries State Funds Appropriated for Research Overhead = Exp. for Admin. and Acad. Supp. Total Investment in Physical Plants	No. of UG Enrolments No. of PG Enrolments Total Semester Credit Hours Generated Amount of Federal Government and Private Research Funds	CCR
Ahn and Seiford (1993)[52]	Examine the sensitivity of DEA, while comparing the efficiency of 80 state and 81 private doctoral-granting universities in the USA for the academic year 1985-86.	Faculty Salaries Physical Investment Overhead Expenses	UG FTEs PG FTEs	CCR, BCC, Additive, Multiplicative
		Faculty Salaries Physical Investment Overhead Expenses	Total FTEs	CCR, BCC, Additive, Multiplicative
		Faculty Salaries Physical Investment Overhead Expenses UG FTEs PG FTEs	UG Degrees PG Degrees Grants	CCR, BCC, Additive, Multiplicative
		Faculty Salaries Physical Investment Overhead Expenses Total Enrolment	Total Degrees Grants	CCR, BCC, Additive, Multiplicative

Comparing Universities

Study	*Purpose*	*Inputs*	*Outputs*	*Model*
Breu and Raab (1994)[53]	Compare the efficiency in the 'production' of student satisfaction and perceived quality of the 'Top 25' Universities, and 'Top 25' Liberal Arts Colleges in the USA in 1992.	SAT/ ACT average or midpoint Percentage faculty with doctorates Faculty to student ratio Educational and general exp. per student Tuition charges per student	Graduation rate Freshman retention rate	CCR
Athanassopoulos and Shale (1997)[54]	Compare cost and value-added efficiencies of universities 45 'old' universities in the UK for the academic year 1992-93.	General academic exp. Research inc.	No. of successful leavers No. of higher degrees awarded Weighted research rating	CCR, BCC
		No. of FTE UGs No. of FTE PGs No. of FTE acad. staff Mean A-level entry score Research inc. Exp. on library and comp. serv.	No. of successful leavers No. of higher degrees awarded Weighted research rating	CCR, BCC with weights restrictions

Comparing Departments

Study	*Purpose*	*Inputs*	*Outputs*	*Model*
Tomkins and Green (1988)[55]	Evaluate the efficiency of 20 UK university departments of accounting for academic year 1984-85, for six different combination of inputs and outputs.	Average FT acad. staff 82/3-84/5 Salaries acad. & res. staff Other exp.	Average no. UGs 82/3-84/5 No. PGR No. PGT Inc. res. contracts Inc. other res. Other income No. of publications 82/3-84/5	CCR

Study	Purpose	Inputs	Outputs	Model
		Comparing Departments		
Beasley (1990)[56]	Compare the overall efficiency of 52 chemistry and 50 physics departments of UK 'old' universities for the academic year 1986-87.	General exp. Equipment exp. Research inc.	No. of UGs No. of PGsT No. of PGsR Research *inc.* If a dept. is rated star If a dept. is rated A+ If a dept. is rated A If a dept. is rated A-	CCR with weights restrictions
Wilkinson (1991)[60]	Compare the efficiency of business and management cost centre of 27 UK universities, and physical sciences academic subject group, comprised of three cost centres, of 46 UK universities, during 1987/88.	Total recurrent exp. from general inc. attributable to research Recurrent exp. specific on res. grants and contracts Total recurrent from general inc. attributable to teaching FT wholly university financed acad. staff No. UGs Mean score of main UG entry qualifications No. PGR No. PGT	First destinations of graduates No. of first degree graduates No. of higher degree graduates Recurrent specific exp. on res. grants and contracts Res. selectivity exercise rating Total continuing educ. courses	Modified CCR
Johnes and Johnes (1993)[43]	Measure the research performance of 36 UK economics departments over the period 1984-88. 192 DEA runs are performed for different input/ output sets, three representative runs are discussed.	Person-months of teaching and research staff Person-months of both teaching and research plus research only staff Research grants per capita	Papers in academic journals Letters in academic journals Authored books Contributions to edited works Papers in 'core' journals Total research grants	CCR
Sinuany-Stern et al (1994)[59]	Compare the relative efficiency of 21 departments of Ben-Gurion University in Israel in 1988.	Operational expenditures Faculty salaries	Grant money Number of publications Number of graduate students Number of credit hours given by the department	CCR

		Comparing Departments		
Study	Purpose	Inputs	Outputs	Model
Beasley (1995)[57]	Determine jointly teaching and research efficiencies of 52 chemistry and 50 physics departments of UK 'old' universities for the academic year 1986-87.	General exp. Equipment exp. Research inc.	No. of UGs No. of PGsT No. of PGsR Research inc. If a dept. is rated star If a dept. is rated A+ If a dept. is rated A If a dept. is rated A-	Modified CCR, with weights restrictions.
Johnes and Johnes (1995)[44]	Measure the research performance of 36 UK economics departments over the period 1984-88. Three representative runs are discussed.	Person-months of teaching and research faculty employed over the 5 year period	Papers and letters in academic journals Papers in 'core' journals	CCR, the weight attached to output 2 must not be less than the weight attached to output 1.
		Person-months of teaching and research faculty employed over the 5 year period Value of external research grants per faculty member over the 5 year period	Papers and letters in academic journals Papers in 'core' journals	CCR, the weight attached to output 2 must not be less than the weight attached to output 1.
		Person-months of teaching and research faculty employed over the 5 year period Per capita time available for research	Papers and letters in academic journals Papers in 'core' journals	CCR, the weight attached to output 2 must not be less than the weight attached to output 1.
Johnes (1995)[45]	DEA is applied to data from the RAE92, to study the relative performance of 60 UK university departments of economics. The analysis throws light on returns to scale and on the characteristics deemed desirable by the peer review panel.	Sum of academic staff on the institution's payroll and academic staff whose salaries are externally funded.	Articles in academic journals.	BCC
		Sum of academic staff on the institution's payroll and academic staff whose salaries are externally funded.	Articles in academic journals. Income from research council grants.	BCC
		Academic staff on the institution's payroll. Academic staff whose salaries are externally funded.	Authored books Articles in academic journals Short works Research council income	BCC

Study	Purpose	Inputs	Outputs	Model
		Comparing Departments		
Olesen and Petersen (1995)[58]	Application of a method that accounts for unknown amounts of noise in the data, in an evaluation of the research activities in 18 business and economic departments of Danish universities, during the period 1975-86.	Full professors Associate/ assistant professors Research fellows	Books Articles published in Danish Articles published in a foreign language Working papers	Modified CCR.
Doyle et al (1996)[48]	Use DEA to model the RAE92 for 85 UK business schools during the period 1988-1992.	Research active staff	Articles in top American journals Articles in top British journals Art. in other refereed journals Art. in professional journals Art. in popular journals Short works Refereed conference Book reviews Other publications Non-refereed conference Res. council money Non-res. council grants Books Edited books Book reviews PhDs Res. council studentship FTE PGs (excl. PhD & studentships) Non-res. council studentships	Modified CCR (Andersen and Petersen, 1993[62]) with weights restrictions.

Student and Staff Appraisal				
Study	*Purpose*	*Inputs*	*Outputs*	*Model*
Doyle and Green (1994)[61]	DEA is applied to examples of student and staff appraisal. It is a applied to the marks of six courses on the full-time MBA program run at Bath 1988-1989; and to 38 staff loads at a university department in the UK in the late 1980s.	Nominal	Marks for six different courses	CCR
		Nominal	Teaching hours No. of projects plus PGs supervised No. of papers published Research grant money awarded	CCR

2.2.1 Comparing universities

Ahn (1987)[49] in his thesis, under the supervision of Charnes and Cooper, used and tested DEA, for the first time, as a potential tool for use in managerial accounting in comparison with other approaches for measuring the efficiency of higher education institutions. DEA was applied to doctoral-granting universities in the USA, where they used labour and capital inputs to produce three outputs: teaching as transmission of knowledge, research as creation of knowledge, and public services to the community. By comparing DEA results with those secured from ratio analysis supplemented by trend analysis, DEA was found to be superior not only in raising questions but also supplying decision aids with accompanying supporting information. The DEA results were also compared to econometric regression studies of educational activities. Departing from these previous studies, institutions were divided into those with and without medical schools. For both groups, contrary to the 'statistical averaging' approach used in the econometric studies, public universities proved more efficient than private universities 'on average', when using DEA. The findings are summarised in Ahn et al (1988)[50]. Ahn et al (1989)[51] then go back to Ahn's (1987)[49] suggestions for future research to extend the study to multiple year data. Both static and one-time portrayals are supplemented by time dependent 'window analysis', for five years of data, as intervals between the receipt of funds and the production of results might be needed. Another suggestion for further study in Ahn (1987)[49], is pursued in Ahn and Seiford (1993)[52], which empirically examined the sensitivity of DEA to changes in models and variable sets, while analysing the relative efficiencies of public and private doctoral-granting institutions of higher education in the USA, for the academic year 1985-86. Four different models (CCR ratio, BCC ratio, additive, and multiplicative), and four different variable sets, with different degrees of disaggregation of inputs and outputs, are used. It is concluded that the DEA results strongly support the hypothesis that state schools are more efficient than private schools, when visible and closely monitored output variables are used for evaluation (e.g. FTE enrolments). The reverse is true, however, when quality-related, but loosely monitored, output variables, such as degrees awarded or external grants (over and above state appropriations), are used as output variables. These findings are particularly important for data availability policy, and what exactly is being measured, which will be further discussed in Chapter 4. As to the choice of models, without exception, the hypothesis test results were consistent across all models and variable aggregation, thus establishing the robustness of DEA.

Breu and Raab (1994)[53] used DEA to measure the relative efficiency of the 'best' 25 US News and World Report-ranked universities. Their results indicated how DEA might be used to measure the relative efficiency of these higher education institutions from commonly available 'performance indicators'. The quality ranking of US News bears an inverse relationship to a ranking implied by the narrow production efficiency criterion DEA. In fact, we argue that the results might have been expected, because different purposes for measuring performance were present. Just as perceived quality indicators are useful in the context of choosing a university, efficiency indicators are more useful in the context of administering these institutions.

In the UK, Athanassopoulos and Shale (1997)[54] examined the cost and value-added efficiencies of 'old' universities, thus recognising that different objectives exist in the context of university performance measurement, which might or might not be in conflict. Unlike the previous US-based studies, absolute freedom of weights is rejected. They use virtual weights restrictions in their models for three purposes. First, so that no factors in the model are ignored. Second, to translate value judgements, such as state imposed policies on funding. Third, to reflect the association between certain inputs and outputs, such as the contribution of one input to more than one output, for instance.

2.2.2 Comparing university departments

Tomkins and Green (1988)[55] used DEA, for the first time, to evaluate the efficiency of university departments. At the time, there was much concern in the UK over the efficiency of universities, and the UGC was, for the first time, in the process of evaluating academic departments. The study was done for departments of accounting, to test the suitability of DEA for such an evaluation, as some government agencies were considering its use. They present results for six models, defined using varied input/ output measures, with good stability of results. They realised that their DEA analysis does not evaluate quality, and conclude that if DEA cannot completely replace judgement, at least would be able to inform it. Another advantage they found is that because 'universities cannot easily find a niche to opt out of the competition' they are particularly suitable DMUs for a DEA assessment, which allows for total freedom of weights.

Taking a different view, Beasley (1990)[56] presented a DEA model for comparing university departments of chemistry and physics for the academic year 1986-87. He highlighted the fact, that contrary to previous mostly data-driven exercises, it is important to consider the conceptually correct data to be used in a DEA model, to deduce what data should be available. This theme will be further discussed in Chapter 4. In a departure from the Tomkins and Green (1988)[55] study, some measures of quality are introduced, and the relative importance of different measures translated with weights restrictions on the virtuals. The problem on how to consider research income is debated, and the opposite view to Tomkins and Green (1988)[55] is taken. Whilst evaluating the success of a department in attracting research income may be important (as considered by Tomkins and Green, 1988[55]), Beasley regarded it as equally, or more important, to evaluate how effective a department is at converting this money input into outputs. He did consider research income also as an output (which has the perverse effect of double counting), but only as a proxy of quantity of publication, in the absence of this information for the disciplines concerned. We argue that this debate has to do with the reason why the assessment is being done in the first place, which will be discussed later in this chapter. In his defence for comparing departments of the same discipline, Beasley argued against the legitimacy of comparing university departments concerned with different disciplines with each other (as for instance in Sinuany-Stern et al, 1994[59]). He also found the comparison of entire universities misleading, as the results obtained may have nothing to do with efficiency, but may be due to the different balance and mix of disciplines present in different universities. It is for this reason that Ahn (1987)[49], Ahn et al (1988)[50], and Ahn and Seiford (1993)[52] considered universities with and without medical schools separately, and Athanassopoulos and Shale (1997)[54] investigated the science/ non-science divide. Unlike Tomkins and Green (1988)[55], Beasley thought that the use of weights restrictions improves the model, by translating the relative importance of input/ output measures, although he is not quite clear to whom.

Using data for UK universities Wilkinson (1991)[60] used a modified version of DEA to compare the performance of one cost centre, and one academic subject group, taking into consideration their 'specialisation'. This entails the elimination of one or more variables in the analysis for a particular DMU or set of DMUs and the re-calculation of the performance profile pertaining to those DMUs. The issue of effectiveness is dealt with by weights restrictions, although, as with Beasley, it is not clear who defines these.

Following the 1989 UFC Research Selectivity Exercise, Johnes and Johnes (1993)[43] measured the research performance of UK economics departments using DEA. They used a data set which included, not only the data available to the UFC panel, but also full bibliographic details for the years 1984-88. Unlike the UFC peer group, which might have imposed upon all departments its own subjective judgement, DEA uses the available quantitative information in a purely objective manner. 192 DEA runs were performed for different input/ output sets, and, as in previous studies, they concluded that the sensitivity of DEA to changes in the input/ output specification was remarkably small. From the 192 runs, two clusters of efficiencies, using cluster analysis, emerged; the distinguishing feature being the inclusion of per capita research grants as an input. They strongly favoured the view that grants are an input and not an output, but do include grants as an output in some models to acknowledge different perceptions of relative research performance by different people. In fact, efficiency scores generated by models with grants as an input are generally weakly correlated with the UFC's research rating, which contrasts with a strong correlation exhibited by efficiencies derived from models in which per capita research grants are an output. One might then be tempted to infer from these results that the UFC peer review group might have implicitly considered grants as an output rather than an input. Other conclusions are: firstly, the lack of relationship between the efficiency scores and size of department, which is in line with the CRS model used; and secondly, there is only limited evidence (statistically significant only for some models) that efficiency scores are lower for departments located outside England.

Johnes and Johnes (1995)[44] continued to use DEA to further investigate the technical efficiency of UK university departments of economics as producers of research. They departed from the 192 DEA runs in their previous paper[43], choosing three different models to analyse. These models included, unlike in the previous paper, one weight restriction, which accounts for the fact that papers in core journals are considered to be of better quality. As before, they found no relationship between measured performance and size for all models. In the first model, the comparison between the efficiency scores reported and the UFC ratings, obtained by informed peer review, indicate that the correspondence is tolerably close. This suggested that a department's performance, as viewed by the peers, is determined largely by the per capita rates of publication in academic journals, especially the core journals. The second model introduces the value of external research grants per faculty member as a second input into the research production function. As a result, the correlation coefficient obtained between the efficiencies scores and the UFC rating becomes significantly lower than in the first model. Again they argued that if peers consider both grants and publications to be indicative of high research productivity, then they are guilty of double counting, for some of the publications would not have been possible without the favourable research environment created by the grants. Finally, in the third model, the per capita annual number of hours spent on each of three activities: UG teaching, PG teaching, and PG research supervision, is subtracted from an arbitrary constant in order to arrive at a measure of time available for research by each faculty member. The results obtained, however, are very similar to those of the first model.

Ultimately, they raised the problem that if all possible inputs into the production process were included in the analysis, then all DMUs would likely appear to be technically efficient. They did not explicitly recognise that what the UFC was measuring is not necessarily the same as their concept of technical efficiency. It will be later argued in this chapter that precisely because of that, one needs to know who is doing the assessment and for what specific purpose. However, as Johnes and Johnes pointed out the successful execution of a DEA requires that appropriate and consistent data is available, which is not always the case.

Following his work of determining overall efficiency scores for departments of the same discipline[56], Beasley (1995)[57] considered the problem of how to split joint inputs and outputs between the teaching and research activities in some optimal way. In doing so, he estimated the efficiency with which a given department carries out each activity. Considering this problem highlights the issue of how to determine efficiencies when resources are shared between different activities, and a non-linear approach to this issue based upon DEA was presented. This approach, while having wider applicability, is useful in the context of university performance measurement, as previous studies have shown there is a need for obtaining separate measures for the two university functions of teaching and research.

Johnes (1995)[45] provided an update for the analysis of Johnes and Johnes (1993)[43] with the RAE data collected by the funding councils in 1992, to study the relative performance of 60 UK university departments of economics. The paper is innovative in two ways: first, the analysis is the first to include the new universities; second, scale efficiency along with technical efficiency is dealt with. Johnes' paper discussed three sets of DEA runs (see Table 2-1). The overall efficiency scores obtained for the first simple model exhibit a weak correlation with the outcome of the funding councils' peer review process. The introduction of income from research council grants as an output in the second model had a dramatic effect on the overall efficiency scores, and markedly improved the correlation between the efficiency scores and the results of the peer review. This analysis confirmed their similar findings for the 1989 UFC review. In the third model, an example was given with disaggregation of inputs and outputs. The correlation between the efficiency scores thus produced and the outcome of the peer review process was somewhat weaker than that achieved in the case of more parsimonious models. For the author, this indicated that some departments deemed to be relatively weak performers by the peer review group performed relatively well as producers of the output types which have been added to the third model. This would provide evidence in support of the popular perception that books and chapters in books count for little in the peer group's assessment of economics departments' research performance. However, they failed to analyse the confounding effect of the disaggregation of the input set. As to the issue of scale, they found some evidence that some of the larger departments in the sample suffered from scale diseconomies even though they appeared to be technically efficient, whereas some departments which were of sub-optimal size were also technically efficient. The latter currently may not be receiving comparable rewards by the funding councils. We shall argue that by using BCC, while advocating that it is a more 'advanced' model than CCR, is misleading. In fact, it represents different assumptions regarding returns to scale, and thus the results of the 1993 study cannot be compared with the 1995 one. It would be interesting to know the results of the 1995 study had the CCR model also been used.

Olesen and Petersen's (1995)[58] application serves to illustrate a model they devised that can provide estimates of the sensitivity of efficiency scores regarding an unknown amount of noise in data. Their conclusions related to the model itself, and not to the evaluation of the departments and policy implications that might have been derived.

On the other hand, Doyle et al (1996)[48] built on the work of Johnes and Johnes (1993)[43] and Johnes (1995)[45] to examine the judgements made by a panel of experts, during the RAE92, about the research rating of UK business schools during 1988-1992. Doyle et al., though, departed from the Johnes' orientation in using DEA in the traditional economic way of estimating production functions for the departments, to use DEA in an operational research approach, as a decision analysis tool. They used policy capture to determine, and critically evaluate, how business schools were judged. They then suggested methods to improve the process of judgement, principally using modified CCR[62] (the unit under evaluation is not included in the reference set, which allows for the ranking of efficient units) as an idealised model whereby the judged institutions judge themselves, consistent with 17 policy constraints. The paper was particularly innovative, in relation to the others above, in that it explicitly addresses the issue of who is doing the assessment, and for what purpose. They compared the panel's apparent decision criteria against the guidelines laid down by the UFC for the conduct of the RAE92. By using DEA to model the judgement of the panel, they suggested that future panels might like to consider it as a decision support tool, which may improve the quality and defensibility of their judgements. To make explicit what would otherwise remain implicit, allows the decision makers to look for unwarranted policies, which may have crept into their judgemental processes. An example of an unwarranted policy to which the panel has fallen prey, according to the paper, is insular mentality: English universities, old-established universities, the panel's own university, and British journals are favoured; as are large institutions. The DEA model, however, simulates a panel consisting of one member from each of the to-be-judged institutions, who is authorised to behave in a totally partisan manner in suggesting which of the performance indicators is important to their institution, and at the same time having to comply with externally imposed, or logically necessary set of policies. In the DEA results, old universities still performed better, and so did the panel institutions to a lesser extent. However, neither English, nor larger departments performed better. They concluded for the superiority of the model of the judge over the judge: 'The modelling process works because, given sufficient examples of judgements to estimate the judge's policies, the 'noise' of inconsistency of the judge may be removed.'

2.2.3 Appraisals of academic staff and students

Doyle and Green (1994)[61] made an experiment of applying DEA to the problem of student and staff appraisal in higher education. The subject is not strictly in the remit of this thesis, however, they do attempt to address some of the questions considered important in this thesis. They not only discussed the technique itself but also discussed the wider implications of having and using such a technique to assist appraisal. Namely, whether to appraise and for what purpose, what to take into account, how to measure it, and finally how to combine it. They conceded, for instance, in the debate whether research income is an input or an output, that different people, at different times ('perhaps with different hats on') will take different views. They did focus, though, mainly on the question of how to use DEA to sensibly combine different appraisal criteria.

2.3 The Approach of this Thesis: Bringing in The Stakeholders Perspective

As seen from the review above, previous DEA studies typically do not explicitly address the existence of different levels of stakeholders interested in the performance of universities. Here, a theoretical framework for the measurement of performance in universities using DEA, considering not only the evaluative side of the exercise but as a tool for the enhancement of performance and organisational change is developed. For this purpose the approach takes into consideration the existence of different universities' stakeholders and their purposes for measuring performance, against a background of what can currently be said about their mission and objectives.

Universities, as with any other organisations, serve a variety of *stakeholders*, defined as individuals or groups of individuals who have a stake or interest in the continued survival and high performance of the organisation (see Mittroff, 1983[63]). The final report[11] by the Joint Performance Indicators Working Group (JPIWG) identifies the customers for the statistics to be published by HESA and their principal needs. The customers they identify include: higher education institutions, Government, funding councils, students, employers and the public at large.

For the HEMS, the repository of the latter JPIWG, the prime purpose is to identify statistics for use by institutions in their internal management and, in addition, to produce sector-level information for use in reporting on the sector as a whole. Hence, we can already identify an attempt to address the issue of performance measurement for different customers. Whether the measures provided are adequate is a matter of debate that will later be addressed (see Chapter 4).

Previous research has traditionally ignored the existence of many of these 'customers' - stakeholders of the higher education sector - when assessing universities' performance. Some studies adopt an economic perspective of labour and capital inputs producing research, teaching and public service outputs. Others take a partial view, such as *The Times Good University Guide*[64], which claims a student perspective. Finally there is a panoply of partial studies that concentrate only on research or teaching across a particular subject, and that are often driven by data availability rather than by a reflection of what should be measured, for which purpose, and for whose interest.

In our approach we shall differentiate between three different levels of stakeholders for the purpose of universities' performance measurement in the UK (Table 2-2). One can argue that Government represents the wider society and that the funding councils implement governmental policy, and hence can be aggregated at the same level: *the State perspective*. At another level, we have the institutions (constituted by departments with its staff and students), which hold a good degree of autonomy, but are also accountable for state funds: *the institutional perspective*. Finally, we consider the potential student who is in the process of choosing a university: *the applicant perspective*.

Table 2-2: Three perspectives in universities performance measurement

Perspectives	*Stakeholders*
State	Employers and society in general
	Government
	Funding councils
Institutional	Institutions
	Departments
	Students
Applicant	Potential students

According to Mitroff (1985)[63], stakeholders are all those interest groups, parties, actors, claimants, and institutions - both internal and external to the corporation - that exert a hold on it. Corporations face complex, messy problems, for which there are often many equally promising solutions and not just one 'best' solution accepted as such by all stakeholders. We can build an analogy, by considering the higher education system as the corporation with the diffuse 'problem' of performance measurement to be solved. The point is that different stakeholders do not generally share the same definition of an organisation's 'problems' and hence, they do not in general share the same 'solutions'. A method is needed that acknowledges different perspectives since it informs us of different options, and ideally works toward a final point of shared commitment to a set of possible solution alternatives. For that reason, having determined three different standpoints, the following questions need investigation:

- What are the mission/objectives of the different stakeholders?
- What is their purpose for measuring performance?
- What measures are necessary to that end?
- What issues emerge from the interaction of the different perspectives?

The different perspectives of the stakeholders result from the existence of different missions and objectives within the sector, and lead to different purposes of performance measurement. Performance measurement is often used in a formative way. Institutions, for instance, measure their performance to establish targets and to benchmark themselves against other performers in order to enhance their performance. It can also have a summative purpose of evaluation, when used for the allocation of funds by the State depending on merit, or in the choice of a university by an applicant.

To serve the different purposes for measuring performance, different criteria will be required. As an example of this diversity and considering student accommodation, students will be interested in good quality convenient accommodation at a reasonable price. Institutions will additionally be concerned about the level of income from university accommodation, whilst the funding councils will primarily be concerned with the overall cost effectiveness and efficiency of the institutions.

We address these questions in the following sections, starting with the external perspectives on university performance: the *applicant* and the *State* perspectives, and finally with the internal vs. external evaluations: the *institutional* perspective.

2.3.1 The Applicant's Perspective

It is generally assumed that the potential student wants to gain a qualification that will provide him/ her with good job and life skills, apart from having an enjoyable extended period of personal and social development.

The potential student will want to select an adequate university, 'a place to live, work, sleep and play for three or four years', to achieve his/her objectives. However, different prospective students will put different emphases in different parameters when choosing a university.

One of the most popular sources of information that attempts to assess universities' performance from the perspective of the potential student is *The Times Good University*

Guide. It claims to give a balanced picture of UK's universities in order to provide reference material for students choosing an undergraduate degree. The volume for 1996[65] rated universities according to ten criteria: entry points, student/ staff ratios, library spending, student accommodation, teaching assessment, firsts awarded, research ratings, and value added. Then it produced an aggregate performance measure for each university, based on equal weights given to the different factors, allowing for the construction of a league table. For the 1997 edition value added was dropped because it was considered to be an unreliable measure at that point. A different weighting scheme was also introduced to supposedly reflect different concerns of the prospective undergraduate: teaching and research assessments were given weighting of 2.5 and 1.5 respectively, compared to 1 for the rest of the indicators.

The *Times* guide defends that students have a right to 'more objective' information than is provided by university publications. This is probably true to some extent. We shall argue, however, that different students will use different criteria when assessing universities and will have different preference structures relating to the criteria used. From this assumption the effectiveness of 'one-suits-all' league tables as an aid to university selection in questioned. For instance, Alomenu (1994)[66], a mature student at Nottingham University, does not think that the information provided is the most appropriate:

'Out of the Times *headings, the only factors that were important to me were student accommodation and future employment prospects. The other factors didn't affect my choice. My choice of Nottingham was based on the type of course and its quality as perceived from entrance literature and the university's prospectus. Further impressions were gained from my interview, from an alternative prospectus and from talking to a student at the campus.'*

In the face of this account two questions arise:

- If the guides, like the *Times'*, are not giving the information required to potential students, what information should be provided?
- Is there an archetypal student, or should we be catering here for different potential students, with different profiles and consequently needing a different set of measures to cover all tastes?

These questions shall now be addressed based upon what is known about student behaviour and consumer behaviour in general. The choice of a university by a prospective student can be considered a form of complex decision-making behaviour, which is classified by consumer behaviourists (Dibb et al, 1994[67]) as an 'extensive decision-making process'. This procedure comes into play when a purchase, here the choice of a university, involves an unfamiliar, expensive and infrequent acquisition. The buyer, in this case the prospective student, uses many criteria to evaluate alternative choices and spends much time seeking information and comparing universities before deciding on the 'purchase'. Students will have already some information in their memory about universities they may choose. If that is not enough to make a decision they will search for additional information. An individual's personal contacts - friends, relatives, associates - are often viewed as credible sources of information, because the student trusts and respects them. Utilising the university dominated sources of information, which include prospectuses and open-days, is another possibility. Students can also obtain information from public sources; for instance, Government reports, news stories, or guides such as the one published by the *Times*. To evaluate the universities, the student establishes certain criteria, by which to compare, rate and rank the different alternatives. However, the greater the quantity of information available, the higher the probability becomes that the student may be overloaded with information, become confused and make a poorer choice. It is for this reason that it is important to improve the quality of the information, to stress the features important to different students, such that they are able to make an informed decision, helping them make a better 'acquisition'. According to consumer behaviourists, three major categories of influences are believed to affect the consumer buying decision process: personal, psychological and social factors. Personal factors such as demographic factors, situational factors and level of involvement are unique to a particular person. Individual characteristics such as age, gender, ethnic origin, income, family life-cycle may influence the choice of a university. Therefore, it is essential to answer the second question above. What measures do different students need to choose a university?

Some studies (Keen and Higgins, 1990[68] and 1992[69]; Roberts and Higgins, 1992[70]; Allen and Higgins, 1994[71]) have concluded that different kinds of students have different criteria, or at least weight the criteria differently, when choosing a university. For instance, comparing *mature* with *traditional* applicants, shows that mature students are more concerned about distance from home, whereas traditional students give greater consideration to the nature of the town/ city in which the institution is based. When it comes to entry qualifications, these play a less important role for mature students, since they are usually faced with a situation where they will be accepted without formal qualifications. Even among traditional applicants it would be expected that the weight of the entry grades will vary. A *high flyer,* who is expecting top A-level grades, will find it easier to be accepted in almost any university; whereas for a *less able* student the A-level grades will be a stronger constraint to finding a place at a university. Differences will also be encountered when comparing *home/ overseas* and *undergraduate/ postgraduate* students.

2.3.2 The State Perspective

The objectives of the Government/ funding councils normally refer to the pursuit of wide participation[4] in higher education at value-for-money for a better educated and trained workforce[72], and to serve the economy more efficiently[5], by achieving greater commercial and industrial relevance[73, 74] in higher education activity. Lately, the introduction of market forces[5] into the higher education system has been vigorously pursued: universities are expected to compete for students, state funds, and other external income.

There is clearly a main evaluative purpose in measuring the efficiency and effectiveness of the use of public funds distributed by the funding councils for teaching and research. These measures provide the funding councils with information to support their operational requirements in the allocation of resources. On the other hand, there is a secondary purpose of informing the institutions delivering higher education, the Government, the funding councils, students, employers and the public at large, of the performance of institutions and higher education in general. This allows the institutions to benchmark themselves against other institutions (a formative purpose of performance measurement); the Government to monitor the pursuit of its policy for the sector; and the potential students to make more informed choices.

The funding councils, in conjunction with HESA, currently collect data for the construction of performance measures at sector (published[13]) and institutional level (not fully published yet) relating to teaching, research, and finance. For each type there is an impressive array of measures that would be cumbersome to list and analyse individually here. Instead some general observation on classes of measures are made.

Macro indicators refer to the nature of applicants and admissions, participation in higher education, description of student population, progression, costs, length of higher education experience, qualifications obtained, first destinations of graduates, impact on general population; and finance statistics for the whole sector. These and institutional financial profiles have already started to be published. Institutional statistics for the teaching activity will comprise: student progression rates, exit qualifications, and employment destinations. Research indicators will comprise: different sorts of research output, research students, and external research income.

The deficiencies of quantitative measures are somewhat addressed by the measuring of quality through the teaching quality assessment and research assessment exercises. The method of peer review used in those exercises ensures that the volume of activity has passed a test of quality. Nevertheless, comparisons are made only within subject areas and relative performance is not valid across different disciplines, leaving inter-institutional comparisons difficult to make. The measures are mostly output measures, and little account is taken of the input side. Value for money, the much cherished governmental objective, is in fact very difficult to assess, since it is hard to assign a value to the different forms of output, and no account is taken to inputs in the sort of measures available. As to value added, there are problems with the comparability of outputs, such as exit qualifications. Some argue that the system of external examiners in place, does not enforce completely the supposed comparability of qualifications across universities. To standardise and ensure comparability between exit qualifications is, in fact, one of the remits of the new Quality Assurance Agency.

Finally, it is doubtful whether the stated purpose of informing other stakeholders of the sector are being met. Students that are in the process of choosing a university will not have information from this source on special needs provision, health care, counselling services, cost of living and accommodation, or availability of places in halls of residence, for instance. When it comes to institutions, care is needed before official management statistics can be converted to relevant measures of performance, since close attention must be paid to their individuality in relation to their missions.

2.3.3 The Institutional Perspective

The universities are usually intuitively categorised on the basis of their age and subject orientation into different groupings: ancient (English and Scottish), civic or redbrick (old and newer, with the two federal universities, London and Wales, sometimes treated separately), technological, and modern or plate glass. Some authors (King, 1970[75]; Dolton and Makepeace, 1982[76]; Tight, 1988[77]) have used multivariate techniques - such as cluster analysis, factor analysis and discriminant analysis - to arrive to typologies of universities. These typologies, although different in detail, have confirmed the existence of five major groups of universities: London, Oxbridge, civic, technological and campus, to which we have added the new universities group (see Table 2-3).

Table 2-3: Universities typology (based in Tight, 1988)

Grouping	*Description*
London	Other than its size , its major distinguishing characteristic is the high percentage of students taking medical courses.
Oxbridge	Though sharing a number of characteristics with the civic group, it is distinctive in its emphasis on languages and other arts subjects, its extensive residential provision, its large proportion of postgraduate students, and in the numbers of them who are pursuing research as opposed to taught courses.
Civic	These universities have an above average number of undergraduates, a significant involvement in medical education, and a below average proportion of income from research grants and contracts.
Technological	These are of average size, with an heavy emphasis on engineering and professional/ vocational subjects, correspondingly low provision of languages and other arts subjects, and a high proportion of male students.
Campus	These institutions are characterised by their concentration on the study of education, social studies, languages and other arts subjects. They have an above average proportion of students in residence, average or below average enrolments, and a low proportion of income derived from research grants and contracts.
New universities	The former polytechnics that chose to become universities following *The Further and Higher Education Act 1992*. They tend to have fewer postgraduate students and in terms of subject split, they have fewer students studying medicine, science and languages, and more in education, engineering and social studies.

Regardless of the fact that each university might have its own mission and objectives, they generally address similar issues. First of all, they will want to attract students; then to develop competencies in them, while placing different emphases on research/ teaching, regional/ national scope, vocational/ academic courses; and finally, to survive as an organisation.

Emergent issues result from the interaction of the different perspectives on universities' performance. Higher education seeks to meet multiple objectives, for instance, good quality teaching and research; and as it pushes to extremes of excellence on one of these objectives, it increases the probability of not meeting the constraints on others. Similarly, the conflict between stakeholders might occur:

> *'The clash between the **Government** conception of the university, and the **universities'** own conception of the university, is actually quite central to contemporary British life.' (Cowen, 1991)*[78]

> *'educational **institutions** will have to learn to understand the motives of **students** just as businesses have to understand their markets.' (Preston et al., 1992)*[79]

On one hand, universities have to meet goals and objectives set outside themselves by their sources of financial support, the funding councils. Regardless of each university's methodology for defining what is good performance and how to monitor it, they will always be accountable to the public through the funding councils. On the other hand, they will have to satisfy the demands of their market - the students, that will demand the use of increased resources. At institutional level, the evaluation that a university makes of its constituent departments, will have to take into account the external evaluation by applicants and the state (see Figure 2-1).

State		*Institutional*		*Applicant*
perspective		*perspective*		*perspective*
external	⇐	internal	⇒	external
evaluation		evaluation		evaluation

Figure 2-1: Two internal/ external interfaces of institutional evaluation

The two external perspectives will affect the internal performance measurement and control system that allows an institution to correct its behaviour and keep in line with its stated missions. Dyson et al (1994)[80] explain this mechanism of performance measurement and control (see Figure 2-2). Each university defines a mission statement according to the direction it wants to follow. This mission statement then needs to be operationalised into more detailed objectives to be achieved. A set of performance measures are defined to monitor the pursuit of the objectives. According to the results obtained targets may be set to reflect priorities emergent from the mission statement. At this point, behavioural responses in the organisation may be required to accomplish the targets, resulting in an organisational change and ultimately leading to an improved direction.

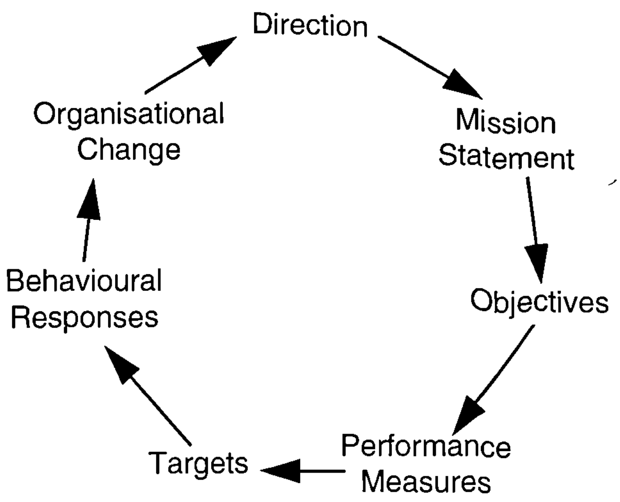

Figure 2-2: Performance Measurement and Control (Dyson et al, 1994)

Through browsing the different institutions' mission statements[81], a large number appear surprisingly similar. This leads one to question whether these are to a large extent for external use as a 'selling' slogan, to please their funding sources and their perceived market.

2.4 Concluding Remarks

The drive for efficiency in the public sector, which started in the eighties, has been extended also to the higher education sector. Universities, as never before, have been subjected to external performance assessment. The drive for assessment came first from the government, who demanded accountability for the resources supplied, and increased efficiency in a time of reducing budgets, and expansion of student numbers. Second, applicants, faced with decreasing generosity from the government in distributing grants, have also increasingly judged their choice of a university more carefully. Universities are jammed in the middle of the external evaluation placed on them, and their own internal assessment for the pursuit of their own objectives.

After surveying the literature on performance measurement in universities, the advantage of using DEA instead, or in conjunction, with PIs, peer assessment, and regression methods is assessed. However, previous DEA analyses of the university sector performance fail, in general, to acknowledge the existence of different levels of stakeholders interested in the measurement of performance. It is ascertained that the interaction of these different levels should be explored. Thus the thesis proposes a framework for the use of DEA to measure performance bringing in the different perspectives on university performance measurement. The application of the framework to the different perspectives is described in the next three chapters.

2.5 References

[1] Cave M, Hanney S, Kogan M, and Trevett G (1988). *The Use of Performance Indicators in Higher Education: A Critical Analysis of Developing Practice.* Jessica Kingsley: London.

[2] Committee of Vice-Chancellors and Principals of the Universities of the United Kingdom, Steering Committee for Efficiency Studies in Universities, Sir Jarratt A (chairman) (1985). *Report of the Steering Committee for Efficiency Studies in Universities.* CVCP: London.

[3] Great Britain, Department of Education and Science (1985). *The Development of Higher Education into the 1990s (Cmnd 9524).* HMSO: London.

[4] Great Britain, Department of Education and Science (1987). *Higher Education: Meeting the Challenge (Cm114).* HMSO: London.

[5] Great Britain, Department for Education and Science (1991). *Higher Education: A New Framework (Cm1541).* HMSO: London.

[6] Committee of Vice-Chancellors and Principals of the Universities of the United Kingdom and Universities Grants Committee (1986). *Performance Indicators in Universities: A First Statement by a Joint CVCP/ UGC Working Group.* CVCP: London.

[7] Committee of Vice-Chancellors and Principals of the Universities of the United Kingdom and Universities Grants Committee (1987). *Performance Indicators in Universities: A Second Statement by a Joint CVCP/ UGC Working Group.* CVCP: London.

[8] Committee of Vice-Chancellors and Principals of the Universities of the United Kingdom (1987). *University Management Statistics and Performance Indicators in Universities.* CVCP: London.

[9] Committee of Vice-Chancellors and Principals of the Universities of the United Kingdom (1995). *University Management Statistics and Performance Indicators in Universities (ninth edition).* Universities Statistical Record: Cheltenham.

[10] Committee of Vice-Chancellors and Principals of the Universities of the United Kingdom (1993). *N/93/51, Research Performance Indicators: Annual Survey of Publications - First Annual Publications Survey: Calendar Year 1991.* CVCP: London.

[11] Higher Education Management Statistics Group (1995). *Higher Education Management Statistics: A Future Strategy.* CVCP: London.

[12] Higher Education Management Statistics Group (1995). *Higher Education Management Statistics: Proposed Publication in 1996.* CVCP: London.

[13] Higher Education Statistics Agency (1996). *Higher Education Management Statistics - Sector Level 1994/95.* HESA: Cheltenham.

[14] Great Britain, Department for Education and Employment (1996). *Higher Education Funding 1997-98 to 1999-2000.* HEFCE: Bristol.

[15] Cave M, Hanney S, Henkel M, and Kogan M (1997). *The Use of Performance Indicators in Higher Education: The Challenge of the Quality Movement. 3rd Ed.* Jessica Kingsley: London.

[16] Universities Grant Committee (1985). *Circular Letter 22/85: Planning for the late 1980s: The Resource Allocation Process.* UGC: London.

[17] Universities Grant Committee (1986). *Circular Letter 4/86: Planning for the late 1980s: Recurrent Grant for 1986/87.* UGC: London.

[18] Universities Funding Council (1989). *Circular Letter 27/89: Research Selectivity Exercise 1989. The Outcome.* UFC: London.

[19] Universities Funding Council (1992). *Circular Letter 26/92.: Research Assessment Exercise 1992: The Outcome.* UFC: London.

[20] Higher Education Funding Council for England (1993). *Circular 7/93. Research Funding Method.* HEFCE: Bristol.

[21] Higher Education Funding Council for England, Scottish Higher Education Funding Council, Higher Education Funding Council for Wales, and Department for Education, Northern Ireland (1994). *Circular RAE96 1/94. 1996 Research Assessment Exercise.* HEFCE: Bristol.

[22] Higher Education Funding Council for England (1997). *Circular 4/97, Funding Method for Research from 1997-98.* HEFCE: Bristol.

[23] Scottish Higher Education Funding Council (1992). *Quality Assessment: The SHEFC Approach.* SHEFC: Edinburgh.

[24] Higher Education Funding Council for England (1993). *Circular 3/93. HEFCE Assessment of the Quality of Higher Education.* HEFCE: Bristol.

[25] Centre for Higher Education Studies, Barnett R (chairman) (1994). *Assessment of the Quality of Higher Education: A Review and an Evaluation.* CHES, Institute of Education, University of London: London.

[26] Higher Education Funding Council for England (1994). *Circular 39/94, The Quality Assessment Method from April 1995.* HEFCE: Bristol.

[27] Higher Education Funding Council for England (1997). *HEFCE Council Briefing No.8.* HEFCE: Bristol.

[28] Smith P and Mayston DJ (1987). Measuring efficiency in the public sector. *Omega* **15**: 243-257.

[29] Smith P (1990). The use of performance indicators in the public sector. *J R Statist Soc A* **153**: 53-72.

[30] Johnes G (1988). Determinants of research output in economic departments in British universities. *Res Policy* **17**: 171-178.

[31] Johnes G (1990). Measures of research output: university departments of economics in the UK, 1984-88. *Econ J* **100**: 556-560.

[32] Johnes J, Taylor J and Francis B (1993). The research performance of UK universities: a statistical analysis of the results of the 1989 research selectivity exercise. *Jl R Statist Soc A* **156**: 271-286.

[33] Taylor J (1994). Measuring research performance in business and management studies in the UK: the 1992 research assessment exercise. *Br J Mgmt* **5**: 275-288.

[34] Taylor J (1995). A statistical analysis of the 1992 research assessment exercise. *Jl R Statist Soc A* **158**: 241-261.

[35] Johnes J and Taylor J (1989). Undergraduate non-completion rates: differences between UK universities. *Higher Educ* **16**: 209-225.

[36] Johnes J (1990). Determinants of student wastage in higher education. *Stud Higher Educ* **15**: 87-99.

[37] Johnes J and Taylor J (1987). Degree quality: an investigation into differences between UK universities. *Higher Educ* **16**: 581-602.

[38] Johnes G, Taylor J and Ferguson G (1987). The employability of new graduates: a study of differences between UK universities. *Appl Econ* **19**: 695-710.

[39] Johnes J and Taylor J (1989). The first destination of new graduates: comparisons between universities. *Appl Econ* **21**: 357-373.

[40] Johnes J (1990). Unit costs: some explanations of the differences between UK universities. *Appl Econ* **22**: 853-862.

[41] Johnes J and Taylor J (1990), *Performance Indicators in Higher Education*. SRHE and Open University Press: Buckingham.

[42] Johnes G (1992). Performance indicators in higher education: a survey of recent work. *Oxf Rev Econ Policy* **8**: 19-34.

[43] Johnes G and Johnes J (1993). Measuring the research of UK economics departments: an application of data envelopment analysis. *Oxf Econ Pap* **45**: 332-347.

[44] Johnes J and Johnes G (1995). Research funding and performance in UK university departments of economics: a frontier analysis. *Econ Educ* **14**: 301-314.

[45] Johnes G (1995). Scale and technical efficiency in the production of economic research. *Appl Econ Lett* **2**: 7-11.

[46] Thanassoulis E, Boussofiane A, Dyson R G (1996). A comparison of data envelopment analysis and ratio analysis as tools for performance assessment. *Omega* **24**: 229-244.

[47] Thanassoulis E (1993). A comparison of regression analysis and data envelopment analysis as alternative methods for performance assessment. *J Opl Res Soc* **44**: 1129-1144.

[48] Doyle JR, Arthurs AJ, Green RH, McAulay L, Pitt MR, Bottomley PA, and Evans W (1996). The judge, the model of the judge, and the model of the judged as judge: analyses of the UK 1992 research assessment exercise data for business management studies. *Omega* **24:** 13-28.

[49] Ahn T (1987). Efficiency and related issues in higher education: a data envelopment analysis approach. PhD thesis, The University of Texas at Austin.

[50] Ahn T, Charnes A, and Cooper WW (1988). Some statistical and DEA evaluations of relative efficiencies of public and private institutions of higher learning. *Socio-Econ Plann Sci* **22:** 259-269.

[51] Ahn T, Arnold V, Charnes A, and Cooper WW (1989). DEA and ratio efficiency analyses for public institutions of higher learning in Texas. *Res Govt Nonprofit Accounting* **5:** 165-185.

[52] Ahn T and Seiford LM (1993). Sensitivity of DEA models and variable sets in a hypothesis test setting: the efficiency of university operations. In: Yuji Ijiri (ed). *Creative and Innovative Approaches to the Science of Management.* Quorum Books: Westport, Connecticut, pp 191-208.

[53] Breu TM and Raab RL (1994). Efficiency and perceived quality of the nation's 'top 25' national universities and national liberal arts colleges: an application of data envelopment analysis to higher education. *Socio-Econ Plann Sci* **29:** 33-45.

[54] Athanassopoulos AD and Shale Estelle (1997). An investigation on the cost and added value efficiencies of higher education institutions in the UK using data envelopment analysis. *Educ Econ J* **5**: 117-134.

[55] Tomkins C and Green R (1988). An experiment in the use of data envelopment analysis for evaluating the efficiency of UK universities departments of accounting. *Financial Accountability Mgmt* **4:** 147-164.

[56] Beasley JE (1990). Comparing university departments. *Omega* **18:** 171-183.

[57] Beasley JE (1995). Determining teaching and research efficiencies. *J Opl Res Soc* **46:** 441-452.

[58] Olesen OB and Petersen NC (1995). Chance constrained efficiency evaluation. *Mgmt Sci* **41:** 442-457.

[59] Sinuany-Stern Z, Mehrez A, and Barboy A (1994). Academic departments efficiency via DEA. *Comput & Ops Res* **21:** 543-556.

[60] Wilkinson RH (1991). The measurement of university performance using concepts derived from DEA. PhD thesis, Stirling University.

[61] Doyle JR and Green RH (1994). Self and peer appraisal in higher education. *Higher Educ* **28:** 241-264.

[62] Andersen P and Petersen NC (1993). A procedure for ranking efficient units in data envelopment analysis. *Mgmt Sci* **39**: 1261-1264.

[63] Mitroff II (1983). *Stakeholders of the Organizational Mind, Toward a New View of Organizational Policy Making.* Jossy-Bass: San Francisco.

[64] O'Leary J (Ed.) (1997). *The Times Good University Guide 1997. Fifth Edition.* Times Books: London.

[65] O'Leary J (Ed.) (1996). *The Times Good University Guide 1996. Fourth Edition.* Times Books: London.

[66] Alomenu G (1994). Fantasy League. *Educ* **184**: 109.

[67] Dibb S, Simkin L, Pride WM and Ferrel OC (1994). *Marketing, Concepts and Strategies, Second Ed.* Houghton Mifflin: Boston.

[68] Keen C and Higgins T (1990). *Young People's Knowledge of Higher Education, the Findings of a Research Programme into the Perceptions of 'Traditional' Applicants.* HEIST/ PCAS: Leeds.

[69] Keen C and Higgins T (1992). *Adult's Knowledge of Higher Education, the Findings of a Research Programme into the Perceptions of Potential Mature Applicants.* HEIST/ PCAS: Leeds.

[70] Roberts D and Higgins T (1992). *Higher Education: The Student Experience, the Findings of a Research Programme into Decision-Making and Consumer Satisfaction.* HEIST/ PCAS: Leeds.

[71] Allen A and Higgins T (1994). *Higher Education: The International Student Experience.* HEIST/ UCAS: Leeds.

[72] Higher Education Funding Council for England (1995). *Promoting Excellence: Annual Report 1994-95.* HEFCE. Bristol.

[73] Great Britain, Office for Science and Technology (1993) *Realising Our Potential: A Strategy for Science, Engineering and Technology. (Cm 2250).* HMSO: London.

[74] Great Britain, Office for Science and Technology (1994) *Competitiveness: Helping Business to Win. (Cm 2563).* HMSO: London.

[75] King J (1970). The typology of universities. *Higher Educ Rev* **2**: 52-61.

[76] Dolton P and Makepeace G (1982). University typology: a contemporary analysis. *Higher Educ Rev* **14**: 33-47.

[77] Tight M (1988). Institutional typologies. *Higher Educ Rev* **20**: 27-51.

[78] Cowen R (1991). The management and evaluation of the entrepreneurial university: the case of England. *Higher Educ Policy* **4**: 9-13.

[79] Preston D, Wilson J and Davies M (1992). Does accommodation play a part in student's choice of university? *Higher Educ Rev* **25**: 57-67.

[80] Dyson RG, Athanassopoulos AD and Thanassoulis E (1994). Performance measurement systems, managerial control and data envelopment analysis. Research Paper 117, Warwick Business School.

[81] Higher Education Funding Council for England (1994), *Profiles of Higher Education Institutions*, HEFCE: Bristol.

3. THE APPLICANT'S PERSPECTIVE

3.1 Introduction

Every year applications from potential students are made through the Universities' and Colleges' Admission Service. Overseas students are also provided with the services of the British Council and UK Council for Overseas Student Affairs. This chapter is concerned with the process of university selection by these applicants.

The appropriateness of the *Times*' league table in guiding applicant's choices is considered, and a case study of applicants at a comprehensive school reported on. The contribution of data envelopment analysis as a decision support technique which can produce customised individual league tables to inform the potential student in his/ her choice is illustrated. The last section explores the design of a decision support system for university selection using DEA with informed judgement.

3.2 The Applicant's Perspective

The starting assumption is that generally the potential student wants to gain a qualification that will provide him/ her with good job and life skills, and also experience an enjoyable extended period of personal and social development.

The potential student will want to select an appropriate university, 'a place to live, work, sleep and play for three or four years', to achieve his/ her objectives. It must be borne in mind, however, that different prospective students will put different emphases on different parameters when choosing a university.

One of the most popular sources of information that attempts to assess universities' performance from the perspective of the potential student is *The Times Good University Guide*[1]. It claims to give a balanced picture of UK universities in order to provide reference material for students choosing an undergraduate course.

It has been argued that different students will use different criteria when assessing universities and will have different preference structures relating to the criteria used. From this assumption the effectiveness of *The Times Good University Guide* (1996) as an aid to university selection is considered, and the contribution of DEA as a more appropriate tool for performance measurement in this instance demonstrated.

3.3 DEA and university selection

The previous section has identified the applicant's interest in the performance of universities as being concerned with criteria related to the choice of university, and the current selection system (1996/7) allows applicants to apply to up to six universities, or more strictly apply for six programmes. *The Times Good University Guide* aims to give a balanced view of British universities and provides information to help applicants make an informed choice. The main feature of the guide is a table ranking the universities in order of merit based on data presented in the book. The data relate to ten performance indicators (Table 3-1) and each university is given a score in the range 1 to 100 on each indicator. The scores are then added together to form an aggregate score, which is then used to form the rankings (Table 3-2). The method of aggregation, in which the scores are given equal weights, assumes equal importance of the criteria, apart from the existence of any major differences in the spread of scores within an indicator, which would implicitly give greater weight to criteria for which the spread of scores is greater. The ten performance indicators are as in Table 3-1:

Table 3-1: The Times 1996 performance indicators

Performance Indicator	*Description*
Entry points	This is a figure for the number of A-level or equivalent points required for entry averaged over the different subject mixes. Where a university takes many students with non-standard qualifications such as access courses, the score will be lowered.
Student/ staff ratios	This measure is the ratio of full-time equivalent students to full-time equivalent teaching staff. Research contract staff and academic related staff are excluded from the calculation. The score for an institution may be biased by subject mix, as some subjects are resourced at a higher level than others.
Library spending	This includes spending on materials and staffing but not on capital costs. The copyright libraries of Oxford and Cambridge receive consequentially high scores although undergraduate students in particular may well not perceive benefits commensurate with the scores. The wide spread of scores for this measure implicitly gives it a greater than equal weight.
Accommodation	This indicates the proportion of full-time students who can be accommodated in facilities owned or controlled by the university.
Teaching assessment	This measure uses information obtained from the teaching assessments although not all subjects have yet been assessed.
Firsts	This is based on the proportion of first class degrees awarded and should be treated with caution as it is mainly controlled by the home institution of the student, although the external examiner system aims to ensure consistency of standards.
Research ratings	This uses information from the Research Assessment Exercise to provide a measure of the quality of research. This measure may be more relevant to postgraduate students.
Value added	This measure is a ratio of an aggregation of completion rates, firsts and employment figures to average entry qualifications. This measure implicitly treats entry qualifications as an input and yields higher scores for those institutions that take in students with lower qualifications and enhance their performance in terms of qualifications and employment.
Overseas students	The score here is the proportion of students from overseas, and will be of particular interest to overseas students and may also be a measure of the international standing of an institution.
Employment	This measure is based on the status of students six months after graduation, and is composed from the percentages in permanent employment, unemployed and going on to research or further study. The proportion in permanent work is weighted most heavily in the overall measure, whilst those in temporary work are excluded, and there is no discrimination between different types of jobs.

It is worth noting that the guide is primarily about first degrees, and that wholly or largely postgraduate institutions such as the London Business School, the Royal College of Art, and Cranfield Institute of Technology are excluded.

Table 3-2: The Times ranking of universities 1996 - top of the table

	entgrad	*staffstud*	*libspend*	*accomm*	*teachass*	*overstud*	*firsts*	*gradempl*	*research*	*valadd*	*Total*	*Rank*
Cambridge	100	93	69	91	100	55	100	86	100	51	845	1
Oxford	98	75	100	86	97	60	66	99	98	60	840	2
Imperial	76	100	44	34	96	65	74	77	97	56	719	3
Edinburgh	72	74	35	40	93	56	45	77	94	76	666	4
StAndrews	73	52	37	100	96	60	40	75	82	51	666	4
LSE	83	42	69	37	98	75	30	70	100	60	666	4
York	74	68	32	79	99	55	44	71	87	52	661	7
Warwick	73	58	28	59	93	55	38	78	94	63	640	8
Bristol	73	73	38	40	88	56	46	81	80	64	640	8
Nottham	75	61	34	43	94	57	37	79	87	69	640	8
UCL	69	89	29	29	93	59	50	67	98	57	640	8
Bath	76	51	34	34	88	57	40	86	90	65	620	12
Manchter	72	65	35	43	93	55	32	74	92	59	620	12
Soton	71	69	33	36	92	55	33	72	79	76	620	12
UMIST	70	62	23	46	77	61	53	74	93	61	620	12
Durham	74	44	26	52	94	59	31	83	84	68	620	12
Newcastle	69	70	37	44	80	58	34	72	79	58	600	17
Lancaster	70	34	34	60	91	56	26	83	89	57	600	17
Birmgham	75	48	33	37	83	55	36	79	80	69	600	17
KingsColl	69	58	35	10	84	60	37	79	84	68	585	20
Glasgow	71	61	34	27	89	58	42	74	68	61	585	20
Sheffield	74	45	29	38	91	55	31	79	80	60	585	20
Sussex	73	41	38	38	77	59	41	69	88	59	585	20

The guide produces a ranking of universities using a single value system (Table 3-2). This may in some sense rank the universities in terms of prestige, but is only one of the many perspectives on universities. For example, a university having a mission with a strong commitment to lifelong learning might consider research ratings, entry qualifications and student accommodation as relatively unimportant measures of performance compared to value added and thus see the rankings based on equal values as largely irrelevant. This diversity of perspective will also be reflected in the attitude of many applicants to the league table, as again different individuals will wish to apply their own values in selecting a short list of universities to apply to, rather than simply accept the top six of *The Times* league table. DEA allows a diversity of weights, and its use in measuring university performance from the applicant's perspective is now explored.

DEA measures the relative performance of organisational units, in this case universities, where there are multiple incommensurate inputs and outputs. In DEA terms *The Times* criteria are outputs that are equally weighted in producing an aggregate performance measure, whilst DEA allows flexibility in the weights used in determining an aggregate measure.

To accommodate the diversity of values of the applicants, DEA was first used to assess the desirability (performance) of universities to six hypothetical applicants. The six model applicants were as in Table 3-3.

Table 3-3: Hypothetical applicants to universities

Applicant	*Description*
Strong	An academic high flyer expecting top A-level grades, giving a high priority to a research environment, the proportion of firsts awarded, teaching quality and high entry qualifications. Typically 18 years old and leaving home for the first time so concerned about the accommodation available. Considers employment prospects, staff/ student ratios, library spending and value added less important.
Less Able	A student who is not so strong academically and is looking for an institution geared to supporting students with lower A-level grades or alternative qualifications. Teaching quality, employment and staff student ratios are of the highest priority, with accommodation also important. The library budget and numbers of firsts may be considered but not weighted heavily.
Mature	An older person who wants to upgrade his/ her qualifications to improve career prospects. Money is likely to be tight ,due to mortgages, and possibly dependants; so library facilities are important. Planning to live at home, so accommodation irrelevant. Teaching quality, employment and staff/ student ratios weighted heavily. Next come firsts and library spending. Research ratings of lesser importance.
Local	A school leaver of average ability who is planning to live at home perhaps for family reasons and possibly financial ones. Employment and teaching quality top priorities, followed by library spending, staff/ student ratios and firsts. Accommodation and research quality not important.
Strong Overseas	Good accommodation provision is very important. Proportion of firsts equally important ,as is entry grades, research and teaching quality, and the proportion of overseas students. A slightly lower priority given to employment, library spending, student/ staff ratios and value added.
Less Able Overseas	Gives highest priority to accommodation, staff student ratios, teaching quality and proportion of overseas students. Firsts and library spending also seen as important. Employment and research quality not a high priority.

To develop appropriate DEA models for each category of applicant, it is necessary to categorise each factor/ performance indicator as either an input or an output, and then convert the applicants' priorities into restrictions on the weights for each factor (for more on weights restrictions, see Allen et al.[2] on absolute weights restrictions, and chapter 6 on virtual weights restrictions). *The Times* guide considers each factor to be an output in DEA terms, and weights them all equally. Categorising entry qualifications as an output implies that the more difficult it is to gain entry to a university, the more desirable to the applicant, but this may be consistent only with the priorities of the strong applicants. It is possible that weaker candidates could still chose to have A-level points demanded by the course they wish to enrol on as an output to reflect the 'reputation' of the course. However, some candidates, specially those with very low grades, might be looking for a more value-added approach. Thus, for the remaining applicants the factor entry qualifications was taken as an input, with the universities seeking to maximise other criteria such as firsts awarded and employment. Hence, the aggregate performance measure now has a value-added orientation, so the specific value-added measure needs to be excluded.

The selection of inputs and outputs for the DEA models and their presumed relative importance is shown in Table 3-4 and Table 3-5.

Table 3-4: Selection of inputs and outputs for the DEA models

	Student Prototypes					
Variable	*Strong*	*Less Able*	*Mature*	*Local*	*LA OS*	*Strong OS*
entgrad	O	I	I	I	I	O
staffstud	O	O	O	O	O	O
libspend	O	O	O	O	O	O
accomm	O	O	-	O	O	O
teachass	O	O	O	O	O	O
overstud	O	-	-	-	O	O
firsts	O	O	O	O	O	O
gradempl	O	O	O	O	O	O
research	O	-	O	O	O	O
valadd	O	-	-	-	-	O
other	Nominal I	-	-	-	-	Nominal I

Table 3-5: Prototypes' preference structures

Prototype	*Preference Structure*						
Strong	accomm, entgrad, firsts, research, teachass	≥	gradempl, libspend, staffstud, valadd	≥	overstud	≥	0.05
Less able	gradempl, staffstud, teachass	≥	accomm	≥	firsts, libspend	≥	0.05
Mature	gradempl, staffstud, teachass	≥	firsts, libspend	≥	research	≥	0.05
Local	gradempl, teachass	≥	firsts, libspend, staffstud	≥	accomm, research	≥	0.05
Overseas (weak)	accomm, staffstud, teachass, overstud	≥	firsts, libspend	≥	gradempl, research	≥	0.05
Overseas (strong)	accomm, entgrad, firsts, research, teachass overstud	≥	gradempl, libspend, staffstud, valadd			≥	0.05

The basic constant returns to scale DEA model with an output orientation is used (see Appendix 3-1). As The Times' data is based on indices and percentages, the use of a VRS model would be more appropriate. Where a nominal input is used the CRS model reduces to VRS, given the output orientation of the models. However, when the range of scale sizes is narrow, as is the range of entry grades here, the difference between CRS and VRS estimates tends to be small. The use of weights restrictions frequently renders VRS models infeasible making CRS the only plausible alternative here, although admittedly, not ideal. The models developed subsequently in the thesis are mostly in terms of volume measures where a CRS assumption is more readily sustainable.

The basic CRS DEA model (without weights restrictions) with the above inputs and outputs determines weights for each university which would show that university in the most favourable light. So, for example, the model would put a very high weight on library spending for Oxford, whilst putting no weight on accommodation for a university with a low proportion of university accommodation. The results would thus be inconsistent with the priorities of many categories of applicant and would therefore not be useful in helping select a shortlist of choices. To incorporate priorities into the models, restrictions were placed on the virtual outputs (the product of the output factor value and its weight), so that the contribution of any factor to the aggregate performance measure could be controlled relative to other factors. The factors were placed into three categories, with the first category factors given higher virtual weights than the second, and the second higher than the third, hence acknowledging the relative importance of different factors (very important, important, less important). A minimum virtual weight of 5% is imposed, in order not to let any university disregard some of the factors in the model. This represents a use of *virtual assurance regions*, which is further discussed in chapter 6. The results of applying the DEA models with the weights constraints are shown in Appendix 3-2.

In each table, the universities are ranked by their DEA score which has a maximum of 100 for the most desirable universities for that category of applicant. The tables also show *The Times'* rank. For the academically strongest students the two sets of rankings are very similar with a rank order correlation coefficient of 0.879 (a coefficient of 1 would indicate perfectly matching ranks). Cambridge which topped *The Times* ranking is again top, and the top ten are almost a perfect match. The rankings of most universities are within ten of each other with the greatest differences including King's College, London which drops from 20th to 49th, equal with Buckingham, which dropped from 30th place.

For the less able students the main difference in the assessments is that entry qualifications are taken as an input, which the universities convert into qualifications and jobs, so there is a value-added orientation to the DEA score and rankings. Robert Gordon's is now ranked at the top, with Luton moving from 96th to 3rd. It is worth noting that the DEA model does not capture length of study as a criterion for choosing a degree course, which could be used as a counter-balance to the low grades demanded by some Scottish universities, like Robert Gordon. However, should an applicant not wish to incur an extra year of study, Scottish universities could be left out of the DEA assessment, or used as an a posteriori criterion to select universities from the table of results. This sort of judgements, that are not modelled directly in DEA will be further discussed later on in this chapter. Many of the 'old' universities such as LSE (93rd from 4th in *The Times* table) and Birmingham (80th from 17th) are unable to compensate for their high entry qualifications although Imperial and UCL remain highly ranked. (For example, the position of LSE was a consequence of having very high entry grades, but a relatively poor staff/ student ratio, whereas Imperial has a favourable staff/ student ratio). The tables are not, however, strongly inversely related and the correlation coefficient of -0.0895 indicates no particular relationship at all. The rankings for mature and local students show a very similar pattern.

For academically strong overseas students, the pattern is similar to the strong UK student, although Buckingham with its high overseas proportion which was given the highest priority, moves from 30th to 21st. The greatest change is Kent, moving from 39th to 25th. For the less able overseas student the rankings are similar to the less able UK students although it should be observed that the less able students (in terms of A levels), whether overseas or not, would find it difficult to gain a place at a university such as Imperial which requires high grades on entry.

The focus on the potential applicant leading to the introduction of diversity in values amongst applicants demonstrates that, although *The Times* league table may be appropriate for the most able students, it is not useful in terms of assisting in the choice of university for other categories of applicants. The DEA approach with its greater flexibility, however, is able to produce league tables more consistent with the requirements of different categories of applicants.

Applicants, of course, do not fall neatly into six categories, and are also interested in subject as well as university information. *The Times* does produce subject league tables but, as with the overall table, they are not appropriate for all categories of applicant. In the next section a system for selection is proposed, which can be tailored to the individual applicant, and includes both subject and university wide information.

3.4 Kenilworth School Case Study

After demonstrating that general league tables such as those produced by *The Times* are only catering for the most academically able students, the aim was to facilitate the design of customised league tables to the individual applicant's preference structure. This part of the research involved working with groups of students at the Castle Sixth Form Centre of Kenilworth School, a comprehensive school with an A level programme of some three hundred students, the majority of whom seek admission to universities across the entire spectrum of institutions. An A level programme of studies comprises two years: Y12 (lower sixth) and Y13 (upper sixth), at the end of which students take national examinations (A levels) in the subjects of their choice.

In order to understand the process of student decision making when choosing a university the following questions need to be answered:

- What criteria do applicants use to assess universities?
- What are the sources they gather information from?
- What relative importance is given to each criterion?
- How is the information aggregated into a meaningful result?

In addressing these questions the studies done by HEIST (Higher Education Information Systems Trust), in collaboration first with PCAS (Polytechnics Central

Admissions System) and after 1992 with UCAS (Universities and Colleges Admissions Service), into the perceptions of traditional[3], mature[4], and international[5] applicants; and university students[6] are relevant.

The case study comprised two parts: the first part looked at the process of decision making retrospectively, making use of Y13 students' experience in applying to university, and the second used that insight to build a model that would help Y12 students in their choice of universities to apply to.

In a first visit to the school a group of Y13 students was invited to discuss the theme, 'How have you chosen the universities you applied to?' Factors that had influenced their choice were discussed, as were the people they had talked to and the prospectuses, guides, newspapers and books they had read. Exhibitions, videos, television programmes and electronic databases were also recalled. The general discussion was followed by brainstorming sessions in small groups to obtain more systematic information. By this process a list of criteria taken into consideration when selecting universities, and a list of sources of information used was obtained (Table 3-6 and Table 3-7).

From the literature and the information gathered in the case study, the list of factors that influence the choice of universities was determined (see Table 3-6). These factors were divided into two groups: *DEA variables* and *non-DEA variables*, depending on the fact if, in principle, they can be modelled or not in a DEA program.

Table 3-6: Factors that influence university selection

DEA variables	*Non-DEA variables*
Entry requirements	Course content
Teaching quality	Teaching methods
Research quality	Course assessment
Pass rates	Reputation/ recommendation
Employment prospects	Sponsorship
Living costs	Town/ city/ area
Housing / accommodation	Distance from home
Age ratios	Campus/ city-based
Gender ratios	Size of university
Race ratios	First impression/ atmosphere
General facilities	
Academic support facilities	
Social facilities	
Sports facilities	

The sources of information used by prospective students (see Table 3-7) can also be divided between the information collected through people, and through a more structured category of purpose-made materials. The latter tend to aggregate quantitative information about the factors influencing applicants' choices.

Table 3-7: Sources of information about universities

People	*Purpose-made materials*
Teachers	CD-ROM databases
Students	Universities prospectuses
Family	Alternative prospectuses
Friends	UCAS books
Careers advisers	Newspapers, videos and television
Careers library	Exhibitions
Open days	The Times Good University Guide

In a second visit to the school the students were asked to answer via a questionnaire, how they rated the importance of the different criteria generated on the previous visit. The results appear in Figure 3-1 (from 1 = not at all important to 5 = very important).

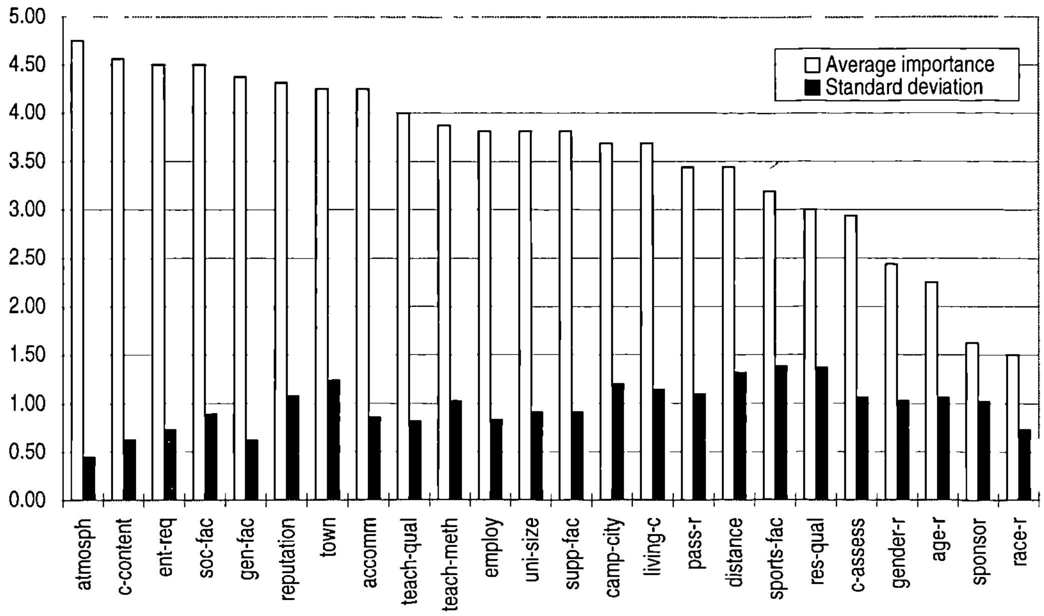

Figure 3-1: Importance attached to different criteria used in university selection

Whereas a lot of the factors that influence university selection are quantitative variables for which information can be easily collected, a lot of them are qualitative variables, which are more difficult to assess and collect information on. However, as can be seen from Figure 3-1, variables such as atmosphere of the institution and reputation are heavily weighted by most students and thus play an important part in their process of choice. This finding is consistent with the fact that many of the published sources of information contain quantitative information, but applicants also talk to a variety of different people in order to assess the softer variables, such as atmosphere and reputation, and are clearly influenced by both kinds of information.

The concept was tested in a second part of the case study, where the DEA models were used as a decision support tool in a university selection workshop, with Y12 students of different academic abilities interested in management and business studies and about to apply to university. The use of DEA for selection follows Doyle and

Green[7], concerning the choice of computer printers, and is effectively using DEA as a multicriteria decision making tool as considered by Belton and Vickers[8], and Stewart[9].

The sources of information used by the students indicated the availability of data, and a match was sought between the available data and the factors mentioned in the first part of the case. From this the inputs and outputs of the model were obtained. The importance given to the different factors were translated into weights restrictions in the DEA models. These different steps in the decision making process and their translation into the DEA models used in the experiment are shown in Table 3-8:

Table 3-8: Comparing universities using DEA

Decision-making		*DEA model*
Sources of information	⇒	Availability of data
Criteria	⇒	Selection of I/O
Relative importance of criteria	⇒	Weights restrictions

- The sources of information used in the process of students' decision-making provide the data that feeds the DEA model.
- The criteria used by the decision maker are either translated into inputs or outputs of the DEA model, depending on factors that the applicant wants to minimise or maximise, respectively. Only factors that are of the type the more the better/ the less the better are modelled in DEA. Other factors have to be dealt with outside the DEA model. Factors such as proportion of firsts awarded and student/ staff ratios, which are used by *The Times*, were not considered, since they were never mentioned by the students in our case study nor in the literature.
- Different preference structures can then be translated into the model by using weights restrictions.

From the list of factors that could be modelled in DEA (see Table 3-6) likely to affect university selection, data was available for all the universities offering programmes in management and business studies on entry grades[1], teaching quality[10], research quality[10], employment prospects[1], cost of living[11], and availability of university accommodation[1]. Data on age[12] and gender ratios[13], and support facilities expenditure[14] was available only for the 'old' universities (i.e. not available for ex-polytechnics and colleges of HE). In Figure 3-1 it can be seen that age and gender ratios were not considered on average particularly important by the students anyway, although some information is lost by dropping the support facilities criterion. It was decided that it was better to use all universities (thus keeping a broader choice) in the models, even though the data available was reduced in that case (thus reducing the number of discriminating factors).

The necessity to drop variables from the prototype model has policy implications with regard to the public availability of data. If the stated purpose of informing all stakeholders on the performance of institutions and higher education in general is to be honoured, the data availability policy needs to be changed. The *University Management Statistics and Performance Indicators*[14] ceased to be published in 1995 (with data concerning the academic year 1992/93) after eight annual volumes had been published. The announced *University Management Statistics - Institutional Level 1995/96*, has still to be published. This means that data on performance is not yet available and that it is not always possible to obtain appropriate information from the raw statistics provided by HESA. Moreover some of the data referred to by students as being important is not provided even in a raw form. A factor such as living costs, for instance, is not addressed in official publications.

In the workshop with the Y12, a pilot was held first with three students to help in the development of the system. A second workshop was then held with eleven students who were first asked to choose whether to consider entry grades as an input (value-added models) or as an output to be maximised according to their expected performance in their A level exams. They then rated the factors to be maximised (outputs) into three orders of importance: very important, important, and less important.

The approach used for the prototypes of students in the previous section would impose two kinds of limitations in the weights flexibility: a minimum virtual weight of 5% for all factors; and very important factors weighted heavier than important ones, and those in turn heavier than less important ones. That approach proved difficult to implement in the workshops due to the large number of weights restrictions that could be required (for example, if five factors were classed very important and five important, 25 restrictions would be necessary and these would be cumbersome to handle particularly when students wished to change their priorities). As a consequence of this the general approach of categorising by importance was retained, but ranges of virtual weights were constructed for each category of factors. This represents the use of simple weights restrictions, further discussed in chapter 6. The less important factors were not allowed to have (virtual) weights bigger than α. Factors considered important would be allowed to have weights in the range α to 2α, and very important ones allowed weights of at least 2α. These are shown in Table 3-9. α was calculated in such a way that the aggregate of the lower bounds on the virtuals determined 50% of the aggregate performance measure retaining 50% for weights flexibility. The use of DEA in this way can be thought of as a game, with the students selecting a space of permissible weights representing their preferences, and the universities using the flexibility of the space to present themselves in the most favourable light.

Table 3-9: Weights restrictions

$L \leq \alpha$	$L =$ virtual weight for a less important factor
$\alpha \leq I \leq 2\alpha$	$I =$ virtual weight for an important factor
$2\alpha \leq V$	$V =$ virtual weight for a very important factor
	$i =$ number of important factors
$where\, \alpha\, i + 2\alpha\, v = 0.50$	$v =$ number of very important factors

As an illustration, one student selected entry grades as an input (value-added approach), teaching quality, employment rates and availability of accommodation as very important factors (virtual weights at least 14.3% from the formula in Table 9); living costs as important (range 7.1% to 14.3%); and research quality as less important (at most 7.1%). The Robert Gordon University had the highest aggregate performance measure. The associated weights were 14.3% for teaching quality and employment, which are at their lower bounds; 50.29% for accommodation, as Robert Gordon has good availability relative to other institutions with low entry grades; 14.01% for living costs, which is close to the top of its range; and 7.1% for research quality. The weights for the other universities at the top of the student's list are presented in Table 3-10.

Table 3-10: Virtual weights for a customised league table

	Top of Table					
Unit	*Rank*	*Teaching*	*Research*	*Employment*	*Living Costs*	*Accommodation*
RGORDON	1	14.3	7.1	14.3	14.01	50.29
ABERTAY	2	14.3	7.1	35.01	14.3	29.29
GLAMORGAN	3	64.3	0	14.3	7.1	14.3
LUTON	4	14.3	7.1	50	14.3	14.3
KINGSTON	5	57.2	7.1	14.3	7.1	14.3
NOTTTRENT	6	64.3	0	14.3	7.1	14.3

The decision-making process is an iterative one, so the workshop adopted an interactive approach. This feature was obvious already in the pilot workshop, where a student first considered entry grades as an input (value-added model), and then as an output of less importance. She realised after analysing the resulting table of the first choice, that her expected A level results would allow her to consider universities with higher entry grades at the bottom of the table. By then selecting entry grades as an output to be maximised, but not so important as to dominate other factors considered more important to her, she obtained a more satisfactory table in relation to her preference structure. This example shows that the positioning of A-level entry requirements as an input or output will depend very much on the candidate at hand, and for the same candidate both situations might be explored. She then made adjustments to her list by striking out universities according to their location. In accordance with this example, the participants in the second workshop were able to rerun the models, changing the choice of inputs/ outputs and/ or the relative importance of the criteria under consideration. They were then asked to comment on the league tables obtained for their suitability in helping them with their choice and how they would further refine them.

In general the students considered the workshop useful. The customised league tables obtained were commented on as 'being a good starting point', a 'rough guide to investigate further', and 'two of the top universities in my list had low entry grades, which is helpful as I don't think I will get high grades'. Some students mentioned that an a priori selection of universities would be useful, as they wished, for example, to consider only universities offering combined business studies and a language. The elimination of some universities in this way would thus simplify the ensuing analysis. As to what further judgements students would use to refine their shortlists, low living costs seems to be an important factor for a lot of students: 'I would leave out universities that were more than 30 miles away from Kenilworth because I can stay at home and it will bring the cost down'. Other considerations mentioned are the programme of study (structure/ content), the location and type of university (distance from home, accessibility, area, type of campus, atmosphere, age distribution), and the facilities provided (sports facilities, quality and price of accommodation).

3.5 Towards a decision support system for university selection

This experiment led to the design of an ideal decision-making process for university selection depicted in Figure 3-2.

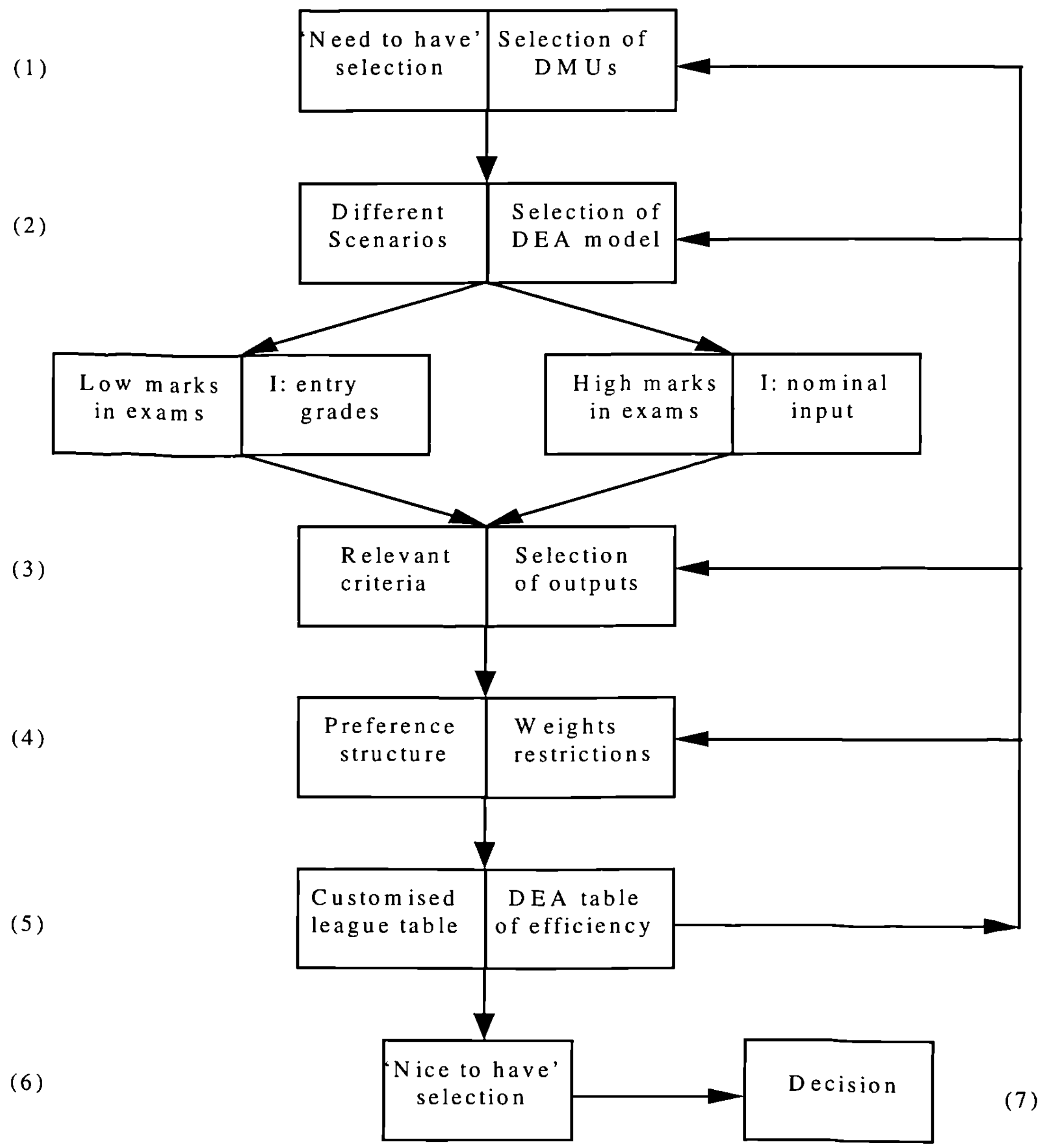

Figure 3-2: Decision-making process for university selection using DEA

(1) **'Need to have' selection:** Here the applicant answers the 'Need to have' question associated with the qualitative variables, which are not modelled in DEA. An initial selection of universities is made by eliminating some based on the course they

want to attend, the distance from home, the location, etc. It corresponds to a selection of units for the DEA model.

(2) **Considering different scenarios:** Depending on the expected results on their exams students will consider different scenarios. The choice will be between a value-added orientation, using entry grades as an input, or an output maximisation only, using a nominal input in the DEA model. Middle-of-the-road candidates might wish to explore both possibilities.

(3) **Choice of relevant criteria:** Here the students will choose which criteria are relevant to them when choosing a university. These will be used as outputs in the DEA model.

(4) **Individual preference structure:** The students will rate the relevant factors in order of importance. These will be translated into the DEA model using weights restrictions.

This task is not a straightforward one as there are several ways to do it, all of them requiring the estimation of appropriate values for the constants in the restrictions, to reflect the value judgements in the efficiency assessments. In this case, the introduction of restrictions incorporating an ordering of the relative importance of the factors to the decision maker was felt to be appropriate. This methodology was introduced in Thompson et al.[15] for absolute weights restriction and Wong and Beasley[16] for proportional virtual weights restrictions. See Appendix 3-1 for the formulations used. The issue of weights restrictions will be further discussed in Chapter 6.

(5) **Customised league table:** The applicants will be provided with a ranked list of universities in accordance with their choices. This corresponds to the DEA table of efficiencies. They can now proceed to the next step or iterate through the process to consider different scenarios, criteria and preference structures.

(6) **Refined league table:** The applicants can now refine the previous table by answering the 'Nice to have' question associated with the qualitative variables. What kind of atmosphere did the university convey during the Open Day? Is it cheap to travel there?, etc.

(7) **Decision:** The final selection is reached by using DEA and informed judgement.

3.6 Conclusion

This chapter focused on the evaluation of universities from the perspective of the prospective student. Applicants to universities differ in their objectives and emphases on different criteria when evaluating universities. DEA is used in an innovative way, allowing for both the institution's and the applicant's views to be considered simultaneously.

The *Times*' league table of universities was considered, and it was shown that a single table weighting criteria equally is not adequate to rank universities given the different missions of the institutions and requirements of prospective applicants. DEA was then introduced as an approach to evaluating universities, which is capable of recognising differences in institutional mission through allowing variability in the weights, and the priorities of applicants by the introduction of weights restrictions. Working with pupils of an A level programme, a prototype system was developed aimed at supporting applicants in producing a shortlist of institutions.

DEA proved to be a useful technique for the aggregation of a confusing mass of information. It can produce customised league tables for individual applicants, and furthermore it can produce several tables, contingent on different preference structures, to help the applicants clarify their ideas. These tables are a valuable aid in reaching an informed decision about the choice of universities to apply to.

As the factors under consideration were principally used as outputs, a CRS output oriented model was used in the analysis, so that the aggregate of the virtual outputs (the product of the output level and optimal weight for that output) summed to 1 (100%) for each unit.

The most appropriate way to implement virtual weights restrictions is not fully resolved. In this study, *virtual assurance regions* and *simple virtual weights restrictions* have been used, as discussed in later in chapter 6.

For the student prototypes' preference structures (see Table 3-5) the following restrictions are added to the model M4 (see Chapter 1, *Introduction*):

$$u_V y_{Vj_0} \geq u_I y_{Ij_0}$$
$$u_I y_{Ij_0} \geq u_L y_{Lj_0}$$
$$u_L y_{Lj_0} \geq 0.05$$

V denotes a very important factor

I denotes an important factor

L denotes a less important factor

In the workshop with the students from Kenilworth School (see Table 3-9) the following weights restrictions were used:

$$u_L y_{Lj_0} \leq \alpha$$
$$\alpha \leq u_I y_{Ij_0} \leq 2\alpha$$
$$2\alpha \leq u_V y_{Vj_0}$$

V denotes a very important factor

I denotes an important factor

L denotes a less important factor

Appendix 3-2: DEA results for the student prototypes

Table 3-11: Strong student model

	Top of Table				*Bottom of Table*		
Unit	*APM*	*DEA Rank*	*Times Rank*	*Unit*	*APM*	*DEA Rank*	*Times Rank*
CAMBRIDGE1	100	1	1	GUILDHALL84	38.04	83	84
OXFORD2	98.63	2	2	WESTENG66	37.96	84	66
STANDREWS4	87.5	3	4	SBANK80	37.74	85	80
IMPERIAL3	86.59	4	3	GLAMORGAN91	37.1	86	91
YORK7	84.2	5	7	CENTENG77	36.91	87	77
EDINBURGH4	81.1	6	4	CALEDONN95	36.37	88	95
LSE4	79.86	7	4	ANGLIA84	36.34	89	84
WARWICK8	79.05	8	8	JMOORES84	36.33	90	84
NOTTHAM8	78.89	9	8	HUDDFLD80	35.29	91	80
UCL8	78.67	10	8	NAPIER80	33.38	92	80
BRISTOL8	78.43	11	8	ELONDON91	32.61	93	91
UMIST12	77.46	12	12	BOURNMTH86	32	94	86
MANCHTER12	77	13	12	THAMESVAL88	31.64	95	88
BATH12	75.6	14	12	LUTON96	31.19	96	96
SOTON12	75.49	15	12	HUMBRSIDE96	28.89	97	96
DURHAM12	74.89	16	12	**Correlation**	**0.879**		

Table 3-12: Less able student model

	Top of Table				*Bottom of Table*		
Unit	*APM*	*DEA rank*	*Times rank*	*UNIT*	*APM*	*DEA rank*	*Times rank*
RGORDON62	100	1	62	EANGLIA32	46.42	83	32
ELONDON91	99.2	2	91	SWANSEA42	46.34	84	42
LUTON96	90.89	3	96	ESSEX29	46	85	29
BUCKHAM30	87.6	4	30	ROYLHOLL24	45.92	86	24
ABERTAY91	82.74	5	91	KENT39	45.22	87	39
DUNDEE30	81.32	6	30	SHEFFIELD20	45.12	88	20
IMPERIAL3	80.3	7	3	SUSSEX20	45.03	89	20
DEMONTFT61	78.08	8	61	HUMBRSIDE96	44.79	90	96
UCL8	76.05	9	8	ASTON32	44.53	91	32
PAISLEY91	73.94	10	91	ULSTER54	43.32	92	54
BRIGHTON66	71.04	11	66	LSE4	43	93	4
OXBRKES57	70.13	12	57	CITY46	42.14	94	46
EDINBURGH4	67.94	13	4	QUEENSBEL47	41.02	95	47
BRISTOL8	66.85	14	8	QMWESTD42	40.07	96	42
NEWCASTLE17	66.68	15	17	KEELE51	39.26	97	51
GLAMORGAN91	65.85	16	91	**Correlation**	**-0.089**		

Table 3-13: Mature student model

	Top of Table				Bottom of Table		
Unit	*APM*	*DEA rank*	*Times rank*	*Unit*	*APM*	*DEA rank*	*Times rank*
ELONDON91	100	1	91	HUMBRSIDE96	47.6	83	96
RGORDON62	100	1	62	SWANSEA42	47.5	84	42
LUTON96	98.63	3	96	ABRYSTTH42	47.3	85	42
DUNDEE30	85.62	4	30	ESSEX29	47.29	86	29
IMPERIAL3	85.32	5	3	DURHAM12	47.22	87	12
BUCKHAM30	84.51	6	30	EXETER35	47.09	88	35
DEMONTFT61	82.3	7	61	QUEENSBEL47	46.49	89	47
UCL8	81.13	8	8	ULSTER54	46.37	90	54
ABERTAY91	78.29	9	91	SALFORD51	46.24	91	51
BRIGHTON66	76.63	10	66	LSE4	45.88	92	4
WESTMSTR66	74.02	11	66	ASTON32	45.71	93	32
NAPIER80	72.8	12	80	HULL47	45.61	94	47
EDINBURGH4	70.55	13	4	KENT39	45.46	95	39
PAISLEY91	69.94	14	91	CITY46	42.86	96	46
BRISTOL8	69.19	15	8	KEELE51	41.06	97	51
NEWCASTLE17	68.54	16	17	**Correlation**	**-0.118**		

Table 3-14: Local student model

	Top of Table				Bottom of Table		
Unit	*APM*	*DEA rank*	*Times rank*	*Unit*	*APM*	*DEA rank*	*Times rank*
RGORDON62	100	1	62	SWANSEA42	49.64	83	42
ELONDON91	97.52	2	91	ABRYSTTH42	49.62	84	42
LUTON96	89.16	3	96	CARDIFF39	49.55	85	39
BUCKHAM30	80.4	4	30	ASTON32	49.16	86	32
ABERTAY91	79.62	5	91	SHEFFIELD20	49.12	87	20
BRIGHTON66	73.69	6	66	KENT39	48.77	88	39
PAISLEY91	72.17	7	91	BOURNMTH86	48.07	89	86
OXBRKES57	71.49	8	57	LSE4	47.93	90	4
DUNDEE30	69.61	9	30	HUDDFLD80	47.93	90	80
WESTMSTR66	68.7	10	66	QMWESTD42	47.3	92	42
KINGSTON57	66.67	11	57	CITY46	45.02	93	46
GLAMORGAN91	66.09	12	91	ULSTER54	44.53	94	54
TEESIDE74	64.38	13	74	QUEENSBEL47	44.4	95	47
IMPERIAL3	63.91	14	3	HUMBRSIDE96	44.29	96	96
NAPIER80	63.52	15	80	KEELE51	42.92	97	51
SUNDLAND77	62.42	16	77	**Correlation**	**0.275**		

Table 3-15: Strong overseas student model

	Top of Table				Bottom of Table		
Unit	*APM*	*DEA rank*	*Times rank*	*Unit*	*APM*	*DEA rank*	*Times rank*
CAMBRIDGE1	100	1	1	GLAMORGAN91	39.82	83	91
OXFORD2	100	1	2	GUILDHALL84	39.67	84	84
IMPERIAL3	90.07	3	3	PAISLEY91	39.48	85	91
LSE4	89.25	4	4	WESTENG66	39.2	86	66
STANDREWS4	87.5	5	4	JMOORES84	38.79	87	84
YORK7	84.22	6	7	CENTENG77	38.71	88	77
EDINBURGH4	81.48	7	4	CALEDONN95	37.73	89	95
BRISTOL8	80.19	8	8	ANGLIA84	37.7	90	84
NOTTHAM8	80.03	9	8	HUDDFLD80	36.93	91	80
WARWICK8	79.18	10	8	NAPIER80	34.63	92	80
NEWCASTLE17	78.3	11	17	ELONDON91	34.49	93	91
UMIST12	78.18	12	12	THAMESVAL88	34.26	94	88
UCL8	78.16	13	8	BOURNMTH86	33.89	95	86
MANCHTER12	78.02	14	12	LUTON96	32.99	96	96
BATH12	77.61	15	12	HUMBRSIDE96	30.42	97	96
SUSSEX20	77.55	16	20	**Correlation**	**0.898**		

Table 3-16: Less able overseas student model

	Top of Table				Bottom of Table		
Unit	*APM*	*DEA rank*	*Times rank*	*Unit*	*APM*	*DEA rank*	*Times rank*
BUCKHAM30	100	1	30	CENTENG77	49.78	83	77
RGORDON62	100	1	62	BRADFORD51	49.76	84	51
ELONDON91	100	1	91	ASTON32	49.7	85	32
LUTON96	83.78	4	96	BIRMGHAM17	49.69	86	17
ABERTAY91	82.72	5	91	SHEFFIELD20	48.92	87	20
DUNDEE30	79.52	6	30	LDSMETRO74	47.07	88	74
IMPERIAL3	78.23	7	3	HUDDFLD80	47	89	80
STANDREWS4	77.81	8	4	JMOORES84	46.43	90	84
PAISLEY91	74.95	9	91	BOURNMTH86	45.66	91	86
UCL8	74.45	10	8	QMWESTD42	45.19	92	42
DEMONTFT61	73.77	11	61	HUMBRSIDE96	44.83	93	96
LAMPETER56	71.53	12	56	CITY46	44.72	94	46
BRIGHTON66	70.71	13	66	KEELE51	43.39	95	51
GLAMORGAN91	68.29	14	91	ULSTER54	41.76	96	54
OXBRKES57	68.29	14	57	QUEENSBEL47	40.73	97	47
YORK7	67.74	16	7	**Correlation**	**0.044**		

3.7 References

[1] O'Leary J (1996). *The Times Good University Guide 1996.* Times Books: London.

[2] Allen R, Athanassopoulos A, Dyson RG and Thanassoulis E (1997). Weights restrictions and value judgments in data envelopment analysis: evolution, development and future directions. *Ann Ops Res* **73**: 13-34.

[3] Keen C and Higgins T (1991). *Young People's Knowledge of Higher Education - Findings of a Research Programme into the Perceptions of "Traditional" Applicants.* HEIST-PCAS: Leeds.

[4] Keen C and Higgins T (1992). *Adults' Knowledge of Higher Education - the Findings of a Research Programme into the Perceptions of Potential Mature Students.* HEIST-PCAS: Leeds.

[5] Allen A and Higgins T (1994). *Higher Education: The International Student Experience.* HEIST/UCAS: Leeds.

[6] Roberts D and Higgins T (1992). *Higher Education: The Student Experience - The Findings of a Research Programme into Student Decision-Making and Consumer Satisfaction.* HEIST-PCAS: Leeds.

[7] Doyle JR and Green RH. (1991). Comparing products using data envelopment analysis. *Omega* **19**: 631-638.

[8] Belton V and Vickers SP (1993). Demystifying DEA - a visual interactive approach based on multiple criteria analysis. *J Opl Res Soc* **44**: 883-896.

[9] Stewart T J (1996). Relationships between data envelopment analysis and multicriteria decision analysis. *J Opl Res Soc* **47**: 654-665.

[10] The Times Higher Education Supplement Internet Service (1996). *Research and Teaching Quality Listings.* Times Newspapers Ltd: London.

[11] Rich J and Footman T (eds) (1995). *The PUSH Guide to Which University 96, 4th ed.* PUSH in association with Lloyds Bank: London.

[12] Universities and Colleges Admissions Service. *Annual Report 1993-94 Entry.* UCAS: Cheltenham.

[13] University Statistical Record (1994). *University Statistics 1993-94.* USR: Cheltenham.

[14] Committee of Vice-Chancellors and Principals of the Universities of the United Kingdom (1995). *University Management Statistics and Performance Indicators.* CVCP: London.

[15] Thompson RG, Singleton FD, Thrall RM and Smith BA (1986). Comparative site evaluations for locating a high-energy physics lab in Texas. *Interfaces* **16**: 35-49.

[16] Wong Y-HB and Beasley JE (1990). Restricting weight flexibility in data envelopment analysis. *J Opl Res Soc* **41**: 829-835.

4. THE STATE PERSPECTIVE

4.1 Introduction

The last years have seen increased pressures for efficiency and effectiveness in the public sector in general, and in the higher education sector in particular. Universities have only been significantly financed by the State since the Second World War. However, since the beginning of the eighties funding cuts have been imposed, more accountability demanded, and selectivity in the allocation of funds, increasingly dependent on the results of performance assessment exercises, has become common policy.

Currently the higher education sector in England funded by the HEFCE (Higher Education Funding Council for England) includes over 200 institutions of higher education - 72 universities, 16 directly funded schools of the University of London, and 50 colleges of higher education. It also funds prescribed courses of higher education at 74 further education colleges, who receive their main public funding from the Further Education Funding Council. There are a further 21 in Scotland, 16 in Wales and 4 in Northern Ireland. In 1995-96 a total of about 1.4 million students followed courses of higher education in UK universities and higher education colleges. Of these about 82,600 were from outside the EU.

Most universities are multi-faculty institutions offering courses in all, or the majority, of the main subject areas. The remaining institutions fall into two broad groups:

- general colleges, offering a range of courses, usually narrower than the universities, and often with a greater emphasis on business and management, or education;
- specialist colleges with more than half their students in one broad subject area.

Universities and colleges receive their income from a variety of sources, public and private. The total amount the funding councils make available each year is determined by Parliament. Funding is provided for teaching and research. The councils decide the basis on which grant is to be distributed and consult the sector on its funding methods. An increased emphasis on quality and performance, and a systematic attempt to assess performance is visible in several official papers:

> *'the general need to contain spending, the pattern of relative costs in higher education, and the demands for capital investment, all mean that a continuing drive for greater efficiency will need to be secured.' (DES, 1991)*[1]
>
> *'The Government believes that the real key to achieving cost effective expansion lies in greater competition for funds and students.' (DES, 1991)*
>
> *'The Council's funding methods are designed to promote stability as far as possible, to maintain quality, deliver efficiency, and to fund research selectivity' and 'is concerned to secure cost-effective use of public funds.' (HEFCE, 1994)*[2]

The growth of the higher education system has raised the question of *effectiveness*, how the actual outcomes compare with the planned aims of the organisation, *efficiency* how resources are deployed to produce different outputs, and *economy*, how the actual resources utilised compare to the planned ones.

Effectiveness is built into the system of performance measurement by translating the parameters defined at governmental level into operational procedures to be carried out by the funding bodies. The efficiency of the system is linked to the assessment of the universities by the funding councils of their teaching, research and financial activities. Finally the economy of the system is regulated by the financial memorandum established between the funding councils and the institutions they fund.

It is the aim of this chapter to explore the use of DEA to support the State and funding councils on, what they claim, is their remit. A critique of how they go about performing their set objectives, specially concerning data collection and availability, is undertaken. However, it is not the objective of the thesis to suggest what the remit of the State and funding councils should be.

In order to become familiar with the work of the funding councils the author has read several past documents published by the council and was part of their mailing list for papers published during the period of writing the thesis. The author has also participated in several meetings organised by the Society for Research into Higher Education where people from the funding councils were speakers and delegates. However, although the work in this thesis has been necessarily informed by contact with the funding councils, the work was in no way commissioned or condoned by them.

4.1.1 The State and Higher Education

The State is responsible for the process of policy formation and implementation. Thus the political parties and Parliament in general, and the DfEE (Department for Education and Employment) in particular, concern themselves with the formation of policy in higher education, and the higher education funding councils (HEFCE, SHEFC, HEFCW, DENI) with its implementation.

Salter and Tapper (1994)[3] point out how the idea of higher education and the relationship between it and the State have changed with time, from more consensual traditions to a more disputed economic ideology, which has created in the recent past a situation of continuing ideological tension. At the beginning of the university era the Christian-Hellenic tradition advocated that the chief duty of the university was to produce good citizens. It should train an elite to be the future leaders in affairs and in the learned professions. Later on a more liberal idea of education emerged, which saw the pursuit of knowledge for its own sake as including the advancement of knowledge as well as its transmission. By 1963, when the Robbins Report[4] was published the traditional ideas of education are still expressed:

- education of 'not mere specialists but rather cultivated men and women';
- the advancement of learning;
- the transmission of a common culture and common standards of citizenship;

but also a more market oriented aspect:

- instruction in skills.

This marks a changing idea of education and will ultimately give rise to a new university model of higher education. The economic ideology of education considers it as an economic resource, which should be organised in a way that maximises its contribution to the UK's industrial development. As a consequence socially relevant, or applied knowledge is more important than pure knowledge. Higher education should be responsive to economic needs, and that it is the responsibility of the State to ensure that universities are held accountable for carrying out their economic role appropriately.

Another change visible in the Robbins Report is the move towards mass higher education, the commitment to the principle of 'social demand': higher education should be available to all those qualified to receive it.

4.1.2 The Current University Model of Higher Education

The current working model of higher education, named by Salter and Tapper[3] a 'managed market', where pedagogical issues are still the concern of the academic faculty but not necessarily matters such as planning, finances and administration.

The new model of higher education in the UK has resulted from the merger of two traditions (former polytechnics and the old universities) under a common university label, which have to continue to live within the framework of new mechanisms of control brought about by the *Further and Higher Education Act 1992*[5]. The new model expressly negates the idea that universities have a common identity and shared purposes; they are designed to be different kinds of institutions performing different missions. The model is a pluralist one, in which different ideas of university are pursued. Their relative success will be dependent upon their ability to secure resources, not only from the State but also the market.

The new model rests on a mass system of higher education, with a purposeful attempt to stimulate demand among previously under-represented groups, which is more responsive to the wider society, and in particular, to the needs of the economy. The transfer of responsibility for the research councils to the Office of Science and Technology in 1992 is another symptom of the ideological change.

Two potential areas for conflict between the State and higher education in the new model will be: the level of resources committed to higher education and the mechanisms that will enable the model to function.

The supply of public funds makes universities amenable to public scrutiny and inevitably they will have to accommodate themselves to public policy determined by the government of the time. Currently the relationship between the State and universities can be described as an attempt on the part of the State to create a managed market financed essentially by public money. The universities have institutional control over their own affairs, while operating within centrally defined and regulated parameters that are managed by the funding agencies.

The funding councils are expected to manage the universities within the framework of ministerial directives designed to ensure that the public receives value for money in return for state funds.

4.1.3 The Future University Model of Higher Education

According to the Dearing Report (NCIHE, 1997)[6], the aim of higher education in the future should be to sustain a learning society. Its four purposes do not differ much from the ones in the Robbins Report, with the economic aspect of higher education being even more emphatic:

- *'to inspire and enable individuals to develop their capabilities to the highest potential levels throughout life, so that they grow intellectually, are well equipped for work, can contribute effectively to society and achieve personal fulfilment;*
- *to increase knowledge and understanding for their own sake and to foster their application to the benefit of the economy and society;*
- *to serve the needs of an adaptable, sustainable, knowledge-based economy at local, regional and national levels;*
- *to play a major role in shaping a democratic, civilised, inclusive society.'*

Since the 1987 White Paper *Higher Education: Meeting the Challenge*[7] the expansion of student numbers has been a characteristic of the UK higher education system. One in five of all 18-19 year-olds in 1992 entered higher education compared with one in seven at the time of the 1987 White Paper. The projection of student numbers in the 1991 White Paper *Higher Education: A New Framework*, indicated that participation rates would continue to increase, and currently already one in three of all 18-19 year-olds enter higher education each year. This resulted in further erosion of the unit of funding for higher education. In 1992 when it was recognised that the response to the call to institutions for expansion had been so successful that participation rates were far ahead of those predicted by the, at the time, Department for Education (DFE) announced that it was going to move to a three year period of consolidation, which has lasted since then. The Dearing Report recommends, however, that widening participation would still be beneficial. It thus advises that students enter into an obligation to make contributions to the cost of their higher education once they are in work, in order to provide resources for growth. Students in this new higher education system will increasingly play the role of 'clients', which should encourage institutions to be more responsive to student requirements:

> *'This will require students to be able to make informed choices based on information about the offerings of higher education, its likely costs, and possible future employment opportunities.' (NCIHE, 1997)*[6]

The issue of data availability in the new 'market' of higher education, where institutions compete for students and resources, will be further discussed in Chapter 5.

4.2 The Mission and Objectives of the Funding Councils

The 1992 *Further and Higher Education Act* put the control of the overall higher education budget, projection of student numbers, the desire for a shift to science and technology, and the steady pressure in favour of the separation of the teaching and research functions into statutory context, reinforcing the role of the government and funding councils. Hence, while the funding councils act on behalf of the universities to government, and may have some policy influence, their main purpose[2] is to perform the following functions:

- the assessment of the quality of higher education;
- the distribution of resources for teaching and research;
- the monitoring of the universities financial probity;

according to guidelines that are controlled at the centre both through legislation as well as by the minister and department.

4.2.1 The Assessment of the Quality of Higher Education

Cave (1988)[8] reports pressures for accountability of the UK higher education system in the same way as other public services that receive money from the Government. In 1985 the Jarratt Report[9], issued by the Committee of Vice-Chancellors and Principals (CVCP) in conjunction with the Government and the Universities Grants Committee

(UGC), suggests that the universities should: work to clear objectives; achieve value for money, and have strong management and planning structures.

The Jarratt Report suggested a major shift in evaluation methods from subjective peer review to performance indicators as a means of performance measurement. It recommended explicitly quantitative as well as qualitative judgements.

In 1986 Government, CVCP and UGC agreed on a Concordat that required changes to be made for further finance to be released to universities. In 1987 the CVCP launched *University Management Statistics and Performance Indicators*,[10] a publication that lists 39 sets of comparative cost data and performance indicators relating to both inputs and outputs in teaching and research for UK universities.

The UGC in 1987 was replaced by the Universities Funding Council (UFC) that became responsible for the distribution of funds among universities under new contract arrangements and a new Polytechnics and Colleges Funding Council (PCFC) was established, which would also submit public sector institutions to a system of contract funding. Performance would be monitored in accordance with those contracts. The Government would thus be concerned to ensure that universities, as much as other public sector institutions, meet objectives determined outside themselves, and demonstrate they achieve these goals. It is in this context that PIs must be considered.

In 1987/88 the provision of Government funds already depended crucially on evidence of real progress in implementing and building upon the proposed changes. There was selectivity in the distribution of resources and rationalisation and, when appropriate, even the closure of small departments as an incentive to better financial management and improved standards of teaching.

Since then systematic evaluation of teaching and research has been undertaken, currently in the form of the Teaching Quality Assessment (TQA) and the Research Assessment Exercise (RAE).

The Teaching Quality Assessment

The higher education funding councils assess the quality of education on a subject by subject basis at higher education institutions on a rolling basis. Assessments are carried out within 61 units of assessment of the quality of education (UÁQE)[11] rather than the broader 11 academic subject categories (ASC) used for funding. The first full cycle of assessments started in 1993 and will be finished by the end of 2001.

The funding councils approach to quality recognises the diversity of institutional mission within the higher education sector, and the quality of the education received by the students is examined within the context of an institution's own aims and objectives. The assessment is by peer review, based on an institutional self-assessment in the subject, with supporting evidence; and on the judgement of the quality of education by a team of assessors during an assessment visit to the unit under review. HEFCE's Teaching Quality Assessment exercise[12] takes into account six aspects of provision when assessing institutions: Curriculum design, content and organisation; Teaching, learning and assessment; Student progression and achievement; Student support and guidance; Learning resources; Quality assurance and enhancement.

When assessing the above, assessors take into account information requested from the subject provider on: entry profile; progression and completion rates; student attainment; employment and further study. This represents a departure from an early outcome based PIs towards a more value-added approach.

Each aspect of provision is rated on a scale of 1 to 4, thus 24 being the highest possible result for a subject being assessed.

The Research Assessment Exercise

The systematic assessment of research started in 1986 by peer review. Since then four RAE's have been performed (1986, 1989, 1992, and 1996). Whereas in the first two exercises all university ('old' universities) departments were assessed, from 1992 onwards institutions were allowed to select which staff and departments to submit. In the RAE96 the ratings were produced for 69 different units of assessment[13] at the same point in time, unlike the teaching quality assessment.

Every time there has been a RAE some modifications to the assessment have been introduced to improve it. Although a peer review exercise, institutions are asked, as with the teaching quality assessment exercise, to provide relevant objective data, which could have a bearing on the judgements to be made. For the last RAE96[14] exercise, the funding councils collected data on the following categories: Overall staff summary; Research active staff details; Publications and other public output; Research students; Research studentships; External research income; Research environment and plans; General observations and additional information. Ratings for the quality of research were awarded in an ascending scale 1, 2, 3b, 3a, 4, 5, 5*.

4.2.2 The Distribution of Resources for Teaching and Research

One of the requirements of the new funding councils after the 1992 *Further and Higher Education Act* was to develop sector wide funding methodologies for allocating resources for teaching and research. These follow with a more or less degree of discrimination the evaluative exercises described above, and are largely formula driven. In addition to council funding, institutions have received undergraduate tuition fees from the DfEE, which mainly flow through LEA's. Under the dual support arrangements, funding council money provides for the cost of the basic infrastructure of university research, other projects being funded by grants from the research councils. Fees for both taught and research postgraduate programmes come principally from the research councils, The British Academy and sponsoring trusts and companies.

Teaching Funding

Teaching funding is a core-plus-margin method with institutions guaranteed a very high percentage in real terms of their previous year's funding, the remaining funds being distributed on a competitive basis. High quality provision will be given an advantage in the competition for the allocation of new funded numbers, low quality provision will be ineligible to compete for additional funded places, and could even have their support withdrawn.

Up to 1996/97[15] funding was distributed through 11 Academic Subject Categories (ASCs), between two modes of study (full-time and sandwich, and part-time) and two levels (undergraduate and postgraduate taught, and postgraduate research). This leads to 44 'cells' for funding purposes. Throughout the method for funding teaching, the key variable is the Average Unit of Council Funding (AUCF). For each institution and each cell in which it is active this is determined by the total funding in the cell divided by the total number of enrolled Home and EU students.

Research Funding

HEFCE allocates nearly all funds for research, as much as 95%, by reference to research volume and quality ratings resulting from the RAE. The remaining funds recognise and encourage generic research, and develop potentially excellent research in institutions not previously funded for research (ex-polytechnics and colleges of higher education). Funds have been attributed in an increasingly selective way, with 75% of the funding provided going to 20 institutions after the results of the RAE92. By far the largest part of volume measure is accounted for by academic staff (around 75%). Up to the 1996/97 academic year the ratings were translated into funding by applying the multiplier Q-1 (where Q is the RAE rating). From 1997/98 funding became more selective, with ratings of 1 and 2 attracting no funding, while a rating of 5* attracts approximately four times as much funding as a rating of 3b for the same volume of research activity.

Proposed Funding

It is proposed that from 1997/98 for research[16], and 1998/99 for teaching[17] to establish four basic levels of resource. These resource levels determine standard prices.

The four groups proposed are Group A: Clinical, Group B: Science, Engineering and Technology, Group C: Other high cost subjects with a studio, laboratory or fieldwork element, and Group D: All other subjects[18].

Accountability for Distributed Funds

For the first time in 1992-93 the grant to be allocated to universities was described by the UFC as 'grant for research' and 'grant for teaching'. Although the money to be distributed remained a block grant, there was a clearer implication that the UFC had expectations, in broad terms, of how the money would be allocated. Although accountability for teaching funds had been in place for some time, via an annual *Redistribution Survey*, the same was not true for research:

> *'At present, for the most part, most institutions are not able to say in detail or in an auditable way what was done with the money allocated by HEFCE'. (HEFCE, 1993)*[19]

By 1993, it was HEFCE's intention to achieve greater accountability also for the use of its research funds, with a new survey, *Accountability for Research Funds.*

Currently, accountability for the funds provided by HEFCE is ensured by three surveys each year. The *Redistribution of HEFCE Funding for Teaching* survey requires institutions to show how they distribute their total grant for teaching. It is one of three surveys required by the Council for funding purposes. This survey and the *Higher Education Students Early Statistics Survey* are required from institutions which receive HEFCE core funds for teaching. The data from these two surveys is then used to inform the funding allocations for teaching in the following year. The data from *The Research Activity Survey* is required only from institutions that expect to receive Council funding for research, and will inform the Council's distribution of research funds for the following academic year.

4.2.3 The Monitoring of the Universities Financial Probity

Besides determining the overall level of its resource commitment, a major concern of the State is to ensure that the institutions act in a financially responsible manner. Hence the funding councils establish a financial memorandum between themselves and the institutions they fund, to monitor the financial health of institutions.

Autonomy from State Funds

On the other hand, the funding councils encourage institutions to exercise their *autonomy* to the maximum degree consistent with full *accountability* for their use of funds provided by the councils[20]. They recognise that institutions obtain funds from other sources, which give them increased scope to pursue their own policies and to take their own initiatives.

4.3 The Contribution of DEA to Inform the State/ Funding Councils in their Remit

In the evaluation of universities by the State, there are different aspects to consider. There is the issue of assessing the quality of provision of teaching and research, and the allocation of resources to the different subject-based units of assessment, which is informed by the quality assessment. The assessment of teaching and research by the funding councils has been based on peer assessment, which is an expensive exercise, and is often accused of being too subjective. The use of DEA to inform peer review exercises to assess the quality of teaching and research can be explored. The single aggregate performance measurement (APM) produced by DEA can be thought of as the equivalent to the rating of such a peer review exercise. Once the assessment of teaching and research is done, the result will inform the allocation of resources. The funding bodies distribute funds for research selectively on the basis of quality. Teaching funds are not selectively distributed on the basis of quality. However, funds are not allocated to units that are deemed not to attain a quality threshold, and when judging bids for increased provision, quality is taken into account.

4.3.1 The Use of DEA to Inform the Assessment of Quality

Peer review assessment is an expensive exercise, in financial cost and human effort. A funding bodies' recent survey put the aggregate cost to the sector of participating in the 1996 RAE at £27.3M. Frequent RAE and TQA are diverting academic time away from research and teaching. Nevertheless the funding councils do need mechanisms which enable them to exercise judgements. These should be robust, widely accepted, and preferably more efficient, effective, and transparent. The funding councils are interested to know the possibilities offered by quantitative methods to assess the performance of universities, specially in relation to research, as seen by the following quotation:

> *'Question 16: If peer review is retained as the primary method of assessment, should this be supplemented by quantitative methods, and if so, how?' (HEFCE, RAE 2/97)*[21]

Doyle et al (1996)[22], as reviewed in Chapter 2, have suggested that future panels might like to consider DEA as a decision support tool, which may improve the quality and defensibility of their judgements. DEA, by making explicit what would otherwise remain implicit, allows the decision makers to look for unwarranted policies, which may have crept into their judgemental processes. To investigate the contribution that DEA may give to answering this question, DEA models of teaching and research activity, which incorporate quality measures, can be envisaged. The results could then be compared to the teaching and research assessment ratings, to determine how they relate.

A DEA Model for the TQA

For the teaching model, and bearing in mind that the quality of teaching provision is assessed by the funding councils on a value-added approach (unlike research), Model 1 is proposed. The model has a value-added approach to quality, as the TQA, with the student entry profile as an input, and completion rates, exit qualifications, and employment destinations, as outputs. FTE academic staff is a measure of input volume, and FTE students, undergraduate and post-graduate taught (post-graduate research students have been considered in the RAE for some time now) give a measure of volume of teaching output for the unit under analysis.

Model 1 considers both quality and volume, which is consistent with the TQA exercise, despite its name. If a pure assessment of quality were required a different model could be envisaged.

Model 1 would be solved for each unit of assessment of the quality of education (UAQE), and the APM obtained comparable to the rating of the TQA exercise. The labels in Model 1 are generic input - output variables only, a detailed discussion of which follows.

Model 1: Teaching Quality Assessment Model

Inputs	*Outputs*
Academic staff	UG and PGT students
Student entry profile	Completion rates
	Exit qualifications
	Employment destinations

Academic staff, measured in FTE, would relate to those academic staff who have teaching activities, and would be a measure of resource volume. *UG and PGT students,* measured in FTE, would represent the volume of teaching activity undertaken. It should be measured by the weighted average of the number of students with different weightings for the mode of study. Information on FTE staff and students for each

UAQE is not available at present, but could easily be available in the submission for the TQA, as it is common for institutions to know the student/ staff ratio for each unit.

Student entry profile represents the quality of the student intake. It could be given by the average A-level points for the students at each UAQE. This information is part of the evidence presented to the assessors at the TQA exercise, although not publicly available. Since the information on entry requirements is available via UCAS for each degree, and again supposing that the mapping to the 61 UAQE is possible, this is a likely measure. The caveat that in some institutions this is not a good measure because of the increasing numbers of non-traditional applicants, whose qualifications cannot readily be compared with A-level applicants, applies also to the TQA exercise.

Completion rates are given by the proportion of those that graduate out of the corresponding cohort of students. *Exit qualifications* are given by the degree classification. *Employment destinations* by the percentage of graduate in employment or further study. Completion rates, exit qualifications, and employment destinations, are all measures of outcome of the teaching process, and as with student entry profile, are part of the evidence presented to the assessors at the TQA exercise, but not made publicly available. They all have some caveats. Low completion rates may identify problems with the applicant selection process, rather than less motivating teaching, and the use of the measure might discourage wider access. Good degree results may reflect high entry scores, but also more lenient standards. The percentage of graduates in employment after graduation, does not capture either long-term employment prospects, or poor 'job matching'. However the combined use of the three measures will tend to balance the picture obtained.

As to the actual measures to be used, there are some problems in using indexes, such as completion rates, in DEA[23]. One alternative would be to use measures of volume moderated by the quality measures. For instance, entry profile, as the volume of students x average entry score. Level of teaching activity, as volume of students x completion rates. As for degree results, the numbers that achieve a certain threshold, e.g. above 2:1 could be used. The same reasoning could be applied to employment destinations.

In conclusion, although the model cannot be solved at the moment, it could in principle be used with existing information, which is unfortunately not publicly available.

A DEA Model for the RAE

The research assessment is concerned only with assessing research quality, it places emphasis on outputs, as opposed to context and research environment. It is a summative process, not overtly concerned with producing formative or developmental outputs, unlike the teaching quality assessment exercise, which has both formative and summative purposes. These characteristics makes it a more likely candidate for replacing it with a quantitative method, than in the teaching peer review process, which produces recommendations for the unit assessed. The exercise informs future funding, by assessing the quality of research conducted in the past. A rating is basically derived based on the quality of published outcomes as well as other research output, produced by those staff submitted for the assessment.

In the light of the considerations above a DEA model of the RAE, Model 2, is proposed. The labels in Model 2 are generic input - output variables only, a discussion of which follows.

Model 2: Research Assessment Exercise Model

Inputs	*Outputs*
Research active staff	Publications and other research output
	PGR students
	External research income

Model 2 should in principle be solved for each UOA, and the APM obtained compared to the RAE rating.

Research active staff, measured in FTE, would be those staff submitted to the RAE. It would be a measure of input volume. *Publications and other research output*, and *postgraduate research students*, would be measures of output volume. Publications and other research output could be divided in sub-categories and thus originate several output measures, for instance, papers in refereed journals, chapters of books, etc. Different importances attached to the different research outputs by the panel of assessors could be translated via weights restrictions as suggested by Doyle et al. (1996)[22] and Johnes and Johnes (1995)[24]. Postgraduate research students, measured in FTE, is another form of research output considered by the RAE.

The issue whether *external research income* should be an input or an output needs clarification. We argue that it should be considered an input when the efficiency of converting this money into research output is being measured (for instance, by the research councils, to assess accountability for their funds). It should be considered an output to measure the success of the DMU in attracting investors, i.e. the competitiveness and reputation (perceived quality) of the unit, which is precisely the case being studied. The idea here, as advocated by Doyle et al[22], is to use DEA as an analysis tool to improve the process of judgement by a RAE panel. The objective is not to devise a model of what should be, as Beasley, 1990[25]; and Johnes and Johnes, 1993[26] propose, which has a concept of technical efficiency behind it. Johnes and Johnes in their analysis of the 1989[24,26] and Johnes in his analysis of the 1993[27] RAE have shown evidence that the panels under consideration have consistently used research grants and contracts as an output (the more the better) in the process of judgement. In the spirit of this thesis, of asking who is doing the assessment and for what purpose, should the evidence of the use of research income explicitly change in future RAE, then its role in the model should change accordingly.

Here it is furthermore argued that *external research income* can be considered a measure of research quality, thus on the output side, as it reflects the prestige of the UOA, and its ability to attract income over and above state appropriations on the input side, which is overtly encouraged by the funding councils. Realising the known caveats of this measure, it would be beneficial to add other measures of research output quality, some of which are presented as evidence by some UOA in their submission to the RAE, such as citation indexes. The number of citations for the outputs of selected researchers, could moderate the volume produced. Another alternative is to use impact factors of place of publication; bibliometric analysis could be undertaken to produce hierarchies of journals within different research areas. As mentioned above, weights restrictions can also be used to translate the relative importance of the output measures to the panel of assessors. Unfortunately the sector has still not agreed on the publication of research performance indicators, that could be used here, although the latter has been planned for some time now.[28]

4.3.2 The Use of DEA to Inform Resource Allocation

As discussed above, the allocation of research money follows the ratings obtained in the RAE, where these are translated into a funding scale, in a more or less arbitrary way. The funding of teaching is not as selective, and the TQA informs on the units that should be funded (provided a threshold of quality is attained) and to inform the allocation of further provision. The use of DEA to help make decisions on funding could be envisaged. The idea is that funding should follow value-for-money for the funding councils, i.e. given the money allocated to the unit, what level of output, both in terms of quantity and quality is achieved. To this purpose, two DEA models are suggested: Model 3 for teaching, and Model 4 for research.

Model 3: Value for Money for Teaching

Inputs	*Outputs*
Redistributed teaching funds	Teaching output

The DMUs for Model 3 would be those candidates in the bid for further student places, the APM for each DMU would represent value-for-money in the teaching activity to the funding council. The resulting top ranking from the solution to the model, would be the shortlist of institutions to receive increased funding.

Redistributed teaching funds here is the way each institution chooses to distribute its teaching grant from the funding council, and it represents the level of resources available to the unit. It is only available by the 11 ASC, not for each of the 61 UAQE, each year as average units of council funding (AUCF). So, if the other measures can be mapped to the 11 ASC, then the model could, in principle, be solved for each ASC.

Since the existence of the teaching quality assessment has not as its main purpose, to inform funding, it will likely remain as a formative exercise, and to inform students of the quality of the education provided by the different units. Thus *teaching output measures* can be calculated directly by measures of volume, e.g. FTE students, moderated by a quality measure, e.g. a teaching quality rating. Otherwise the quality variables for both input and output sides as in Model 1 should be used.

Model 4: Value for Money for Research

Inputs	*Outputs*
Redistributed research funds	Research output

As to research, Model 4 would give a direct measure to be used in funding, instead of the current procedure of translating the research quality rating into a somewhat arbitrary funding scale. Another obvious advantage is to fund according to value-for-money (value-added approach) rather than on outcome basis (as at present), that take into account the conditions the research unit faces (the reason Beasley[25], and Johnes and Johnes[26] give to use research income on the input side). Units that perform well relative to their input levels, currently might not see their performance rewarded in the funding mechanism. Whereas with this value-added DEA approach, they would receive 'fair' funding, which would allow them to produce outputs accordingly.

Redistributed research funds is the only input in the model. As to the internal redistribution of funding council research funds, the first question concerns the units or subject divisions used for reporting on the allocation of research grant to the funding council. A range of options including the units of allocation used to allocate the research grant (72 UOA in RAE92, 96 UOA in RAE96), the cost centres (35 up to 1996/97, 41 since), or the 11 academic subject categories (ASC) used in allocating funding for teaching. Up to the academic year 1996/97, HEFCE has concluded that institutions should be required to report on their allocation of research funding using their own allocation units[29]. This resulted in a range of approaches to reporting and summation of the results that was difficult for comparison purposes. HEFCE therefore imposed a greater degree of standardisation[30]. Institutions are from academic year 1997/98 required to account for the distribution of their research funds by cost centre. Central costs, such as for libraries, computing services and administration, may be shown in the non-departmental cost centres or split out proportionally to the funds in each academic cost centre. When HEFCE collates the data, it allocates central costs to academic cost centres. This means that the required data will be available by academic cost centre, and since HEFCE provides a mapping of UOA and ASC to ACC, the model could , in principle, be solved for each ACC.

Research output measures would be as in Model 2, or if the outcome-based research assessment exercise were kept (either in its current or DEA version) to inform the sector at large on the quality of research undertaken in different research units, research output measures could be calculated directly by measures of volume, e.g. FTE research active staff, moderated by a quality measure, e.g. a research quality rating.

4.3.3 The Link Between Quality Assessment and Resource Allocation

A theoretical framework (see Table 4-1) for using DEA for the assessment of teaching, which can then inform the allocation of resources can now be summarised:

Assessment of the quality of teaching:

1. Selection of DMUs for Model 1, is based on the choice of the subject based units for the assessment of the quality of education defined by the funding councils. It is not always clear to which unit a department should be submitted to.

2. The relevant criteria should be decided by the panel of assessors. This will correspond to the inputs and outputs in Model 1. If the relative importance of criteria can be established by the panel (e.g. firsts awarded > other classifications, etc.), these should be translated by weights restrictions. The final evaluative model is obtained.

3. A quality teaching rating is obtained that corresponds to the APM obtained by solving Model 1.

4. Some moderation might be necessary to accommodate factors that have not been included in the model. And/ or iterations in the preceding steps are deemed necessary to

arrive to the final teaching quality rating. These will inform applicants to university on the quality of provision by different departments, and funding by the funding councils.

Table 4-1: Assessment of the Quality of Teaching and Allocation of Resources for Teaching

Assessment of Teaching Quality				*Funding of Teaching*			
1. Choice of UAQT	2. Choice of relevant criteria, and their relative importance	3. Teaching quality rating	4. Moderation: final rating	5. Choice of units to consider for further provision funding	6. Choice of relevant criteria, and their relative importance	7. Ranking of efficiency in the provision of teaching	8. Moderation: final choice of units to receive further funding
Model 1: Teaching Quality Assessment Model				Model 3: Value for Money for Teaching			
• Selection of DMUs	• I/O choice, and weights restrictions	• APM		• Selection of DMUs	• I/O choice, and weights restrictions	• APM	

Table 4-2:Assessment of the Quality of Research and Allocation of Resources for Research

Assessment of Research Quality				*Funding of Research*			
1. Choice of UOA	2. Choice of relevant criteria, and their relative importance	3. Research quality rating	4. Moderation: final rating	5. Choice of units to consider for research funding	6. Choice of relevant criteria, and their relative importance	7. Ranking of efficiency in the provision of research	8. Moderation: final choice of units to receive further funding
Model 2: Research Assessment Exercise Model				Model 4: Value for Money for Research			
• Selection of DMUs	• I/O choice, and weights restrictions	• APM		• Selection of DMUs	• I/O choice, and weights restrictions	• APM	

Funding for additional places:

1. The funding councils might, based on the teaching quality rating previously obtained, bar units that do not achieve a certain threshold to be considered for the provision of further student places for the subject under consideration. The ones that bid for additional provision, and are suitable candidates, will be the DMUs for Model 3.

2. The relevant criteria should be decided by the funding councils. This will correspond to the inputs and outputs in Model 3. If the relative importance of criteria can be established by the panel (e.g. students from ethnic minorities > other students, if wider access is the objective, etc.), these should be translated by weights restrictions. The final value-for-money model is obtained.

3. A ranking of institutions is obtained in terms of value-for-money to the funding council. The institutions at the top of the shortlist should be receiving the further provision.

4. Some moderation might be necessary to accommodate factors that have not been included in the model, for instance, issues of location of provision around the country. And/ or iterations in the preceding steps are deemed necessary to arrive to the final decision.

In the same way a theoretical framework (see Table 4-2) using DEA for the assessment of research, which then informs the allocation of resources can now be summarised:

Assessment of the quality of research:

1. Selection of DMUs for Model 2, is based on the choice of the UOA defined by the funding councils. It is not always clear to which unit a department should be submitted to, and this step might need iteration, as the panel of assessors might wish to refer it to another UOA panel.

2. The appropriate criteria, that might change for different research areas, should be decided by the panel of assessors. This will correspond to the inputs and outputs in Model 2. The relative quality of outputs will be established by the panel, and these should be translated by weights restrictions. The final evaluative model is obtained.

3. A quality research rating is obtained that corresponds to the APM obtained by solving Model 2.

4. Some moderation might be necessary to accommodate factors that have not been included in the model. And/ or iterations in the preceding steps are deemed necessary to arrive to the final teaching quality rating. These will inform applicants to university on the quality of provision by different departments, and funding by the funding councils.

Selective funding for research:

5. The funding councils might, based on the research quality rating previously obtained, consider just the top of the league for research funding (as it currently does), or keep all the units in and make the decision on which ones to fund on the value-for-money ranking instead. The units considered for funding will be the DMUs for Model 4.

6. The relevant criteria should be decided by the funding councils. This will correspond to the inputs and outputs in Model 4. The relative importance of criteria should be translated by weights restrictions. The final value-for-money model for research is obtained.

7. A ranking of institutions is obtained in terms of value-for-money to the funding council.

8. Some moderation might be necessary to accommodate factors that have not been included in the model. And/ or iterations in the preceding steps are deemed necessary to arrive to the final decision. Institutions can then be funded proportionally to their

efficiency score, or again the funding councils might choose to fund only institutions that achieve a minimum efficiency score.

In conclusion, two purposes for measuring performance at subject level by the funding councils are presented. The first assesses the quality of teaching and research to inform the 'clients' of higher education, the second assesses the efficiency in the use of past resources to inform the allocation of future funds. The assessment of quality informs to different extents the allocation of resources for teaching and research.

4.4 The Performance of the University Sector

Once resources have been allocated to universities, based on the results of the assessment of teaching (mostly in the long run) and research (mostly in the short and medium run), how are these used? On the one hand, there is the issue of accountability for research and teaching funds allocated to universities; on the other hand the councils' encouragement of autonomy of universities from state funds. What is the overall performance of universities in achieving simultaneously these two objectives?

The efficiency of the use of resources, either from the state or not, is also of interest. In fact, when the funding councils allocate extra student places, they take into consideration unit costs for the units applying, in order to assure value for money.

Finally, the effect of subject mix in the measurement of performance of universities will also be considered.

Again, DEA models that take into account both teaching and research activities can be envisaged to address these issues. The results will elucidate such policy questions, as the difference of institutional performance between the 'old' and 'new' universities. Is it fair to assess both types using the same premises, as they were for so long managed and resourced in different ways? Are their performances really different? There is also evidence in previous studies (Johnes and Taylor, 1990)[31], that, at least for the 'old' universities there were differences of performance between English and peripheral (non-English) universities. Is that still true for the new enlarged sector? There is also the issue of Oxbridge performance. Given their preferential treatment in terms of funding for tuition, do they consequentially deliver more value added accordingly? Or, as some contest (THESIS, 1997)[32] the money would be better used if re-allocated to other institutions?

4.4.1 Accountability for and Autonomy from State Funds

In the State/ funding councils perspective, good performance is not only to provide good quality teaching and research, at value for money, but also to ensure some degree of autonomy and financial health. Basically, a comprehensive performance model will be useful, that takes both accountability to state funds and desired autonomy of the institution into consideration. This model would measure the effectiveness of state investment, in triggering off external investment on universities, without leaving out the efficiency of universities in the use of state funds. This is translated in Model 5, where *state appropriations* are the input to the model, *teaching* and *research* are outputs, and so is *earned income*, as a measure of autonomy. The resulting APM is as a measure of overall performance of the university from the state perspective. The model would be solved for each university.

Model 5: Overall Performance

Inputs	*Outputs*
State appropriations	Teaching output
	Research output
	Earned income

State appropriations ideally would be in the form of funding councils grants plus mandatory academic fees channelled through the LEAs. *Earned income* would be any income over and above the state appropriations, such as full-cost fees, research grants and contracts. *Teaching output* could be obtained by summing a measure of volume, e.g. FTE students, moderated by a measure of quality, e.g. the teaching quality rating, for all UAQE across the institution. *Research output* could also be obtained by summing a weighted sum of measures of volume provided in the RAE (FTE research active staff, research assistants and fellows, and research students), multiplied by a measure of quality, e.g. RAE rating, for all UOA across the institution.

4.4.2 Efficiency in the Use of Resources

If the State/ funding councils perspective is considered, the main objective is to provide good quality teaching and research at value for money. Thus when considering future funding, specially when broadening provision, the funding councils look for universities that use resources more efficiently, to award extra provision (for instance, in the allocation of more student places). The efficiency of a university's operation is translated in Model 6, where *total resources* are the input to the model, and *teaching* and *research* are outputs. The resulting APM is as a measure of efficiency (value for money) in the use of resources. The model would be solved for each university.

Model 6: Efficiency

Inputs	*Outputs*
Total resources	Teaching output
	Research output

4.4.3 Disentangling the Subject Mix Factor from the Measurement of Performance

In the above two models (Model 5 and Model 6), the subject mix factor is not considered. However there is evidence from previous studies[31], that this is an important factor. Handling subject mix in the context of comparing universities' performance has been a problem. It is proposed the use of weights restrictions that link the input-output divide to deal with subject-mix effects in this context. The procedure used in this chapter follows the method proposed by Thanassoulis et al[33] to link factors across the input-output divide. It prevents units from taking undue advantage of weight flexibility irrespective of the established links between certain inputs and outputs. The method is applied to the present problem of taking account of subject mix in the measurement of performance. To be able to disentangle the subject-mix factor in the measurement of performance Model 7 and Model 8, corresponding to Model 6 and Model 5 respectively, are proposed. Inputs and outputs are disaggregated into the summed values for different subject cost bands, so they can be linked through weights restrictions, as will be shown in the next section

Model 7: Overall Performance Taking Account of Subject Mix

Inputs	*Outputs*
State appropriations for cost band X	Teaching output for cost band X
	Research output for cost band X
	Earned income for cost band X

Model 8: Efficiency Taking Account of Subject Mix

Inputs	*Outputs*
Total resources for cost band X	Teaching output for cost band X
	Research output for cost band X

Traditionally, researchers have considered a science/ non-science divide when assessing the subject mix on university performance. However, HEFCE proposes[16, 17] four empirically derived cost bands: *Clinical; Science, engineering and technology*; *Other high cost subjects with a studio, laboratory or fieldwork element*; and *All other subjects*. Considering these bands, the number of inputs would raise to four, and the number of outputs from three to twelve! These might lead to lack of discrimination between the units being assessed, which the use of weights restrictions to link outputs to inputs might counteract.

In the next section, an application to the university sector for the academic year 1995/96 to test the usefulness of the proposed models, is described.

4.5 An Application to the University Sector for the Academic Year 1995/96

As discussed above neither Model 1 nor Model 2 can be solved at the moment, because, although data is collected for the inputs and outputs considered, not all is in the public domain. Model 7 cannot be solved either, because data on earned income by cost band is not collected, thus not available. Model 5, Model 6 and Model 8 can however be solved, despite some caveats, that will be discussed. The data used can be found in Appendix 4-1.

4.5.1 Units of Assessment

The unit of assessment for Model 5, Model 6 and Model 8 is the university. Out of the 96 universities (eight of the London University colleges are directly funded by HEFCE, and thus are treated as independent institutions) for which *The Times* compiles the average teaching and research ratings, seven were excluded for different reasons. University of London, Goldsmiths' College concentrates its activities in cost bands B and C; University of London, Imperial College concentrates its activities in group B, and has very little activity in groups C and D; London Guildhall University is the only institution that has zero research activity for some of the groups where it is active in teaching; University of London, LSE is mainly active in group D, with only some activity in group C; UMIST is mainly active in group B, with some activity in groups C and D. University of Wales, Lampeter is the smallest of all institutions and is mainly active in group D, with only some activity in group C; University of London, Queen Mary and Westfield College dedicates more than half of its resources to group A. From *The Times* list three universities were already excluded: the Open University for having very different characteristics from the other institutions; Cranfield University for being mainly a postgraduate institution; and Buckingham, which does not receive funds from HEFCE, and therefore does not have the quality of its teaching and research assessed.

The exclusion of these institutions guarantees a well-conditioned data matrix. This will be specially important in the use of weights restrictions in Model 8 to link inputs to outputs in the same cost band. The models have therefore been solved for 89 universities.

All data used is available from HESA reference volumes *Students in Higher Education Institutions*[34] and *Resources in Higher Education Institutions*[35], with the exception of average teaching and research ratings obtained from the *Times Good University Guide*.[36]

4.5.2 Experimental Design

Model 5: Overall Performance was subdivided into two models with different degrees of disaggregation, as in Table 4-3.

Table 4-3: Model 5 Specification

	Variable Sets	
Model	*Inputs*	*Outputs*
51	HEFCS_INC	NON_HEFCES_INC
		TEACHING
		RESEARCH
52	HEFCS_INC	NON_HEFCS_INC
		TEACH_UGS
		TEACH_PGS
		RESEARCH

The variables used are defined as follows:

HEFCS_INC: *Funding Council Grants* income in £ thousands. It is directly available from HESA statistics. It includes institutional funds for academic and other purposes allocated primarily by the Higher Education Funding Councils for England, Wales, and Scotland, and by the Department of Education for Northern Ireland acting as a funding agency for the two Northern Ireland universities.

NON_HEFCS_INC: sum of *Academic Fees and Support Grants*, *Research Grants and Contracts*, *Other Operating Income*, and *Endowment Income and Interest Receivable*, in £ thousands. All elements are directly available from HESA statistics. *Academic Fees and Support Grants* includes all income received in respect of fees for students on courses for which fees are charged. *Research Grants and Contracts* includes all income in respect of externally sponsored research carried out by the institution for which directly related expenditure has been incurred. *Other Services Rendered* includes all income in respect of services rendered to outside bodies, including the supply of goods and consultancies.

TEACHING: is the total number of FTE students, as the measure of volume, multiplied by the average teaching rating for the university, as a measure of quality. FTE students by university is directly available from HESA statistics. Teaching rating uses the average score for each university in the funding council assessment completed by the end of 1996. Not all subjects have yet been assessed, but those covering a majority of students now have. In principle, the teaching output should be calculated for every UAQE, and the sum for all UAQE by university obtained. However, as not all units have been assessed and FTE students are not available by UAQE, the average for each university is used instead.

TEACH_UGS: is the total number of FTE undergraduates, as the measure of volume, multiplied by the average teaching rating for the university, as the measure of quality.

TEACH_PGS: is the total number of FTE postgraduates, as the measure of volume, multiplied by the average teaching rating for the university, as the measure of quality.

RESEARCH: is the total number of FTE academic staff, as the measure of volume, multiplied by the average research rating for the university, as measure of quality. FTE academic staff is deemed the most important measure of volume by HEFCE, and was thus chosen. Other measures of volume (e.g. research assistants and fellows, postgraduate researchers) used by HEFCE are not readily available by institution. FTE academic staff is not directly available from HESA statistics. A weighted average of FT+0.5PT was used instead. Research rating is based on the average score for each university in the RAE96 carried out by the three funding councils. The averages are calculated for all university academic staff, rather than only those entered for the assessment.

Model 6: Efficiency was subdivided into eight models with different degrees of disaggregation both on the input and output sides, to test the robustness of the results. Despite total income and total expenditure by university having a correlation coefficient of 1.00; there are two 'families' of models: Model 61x and Model 62x with income and expenditure variables, respectively, on the input side; again to test for robustness.

Table 4-4: Model 6 Specification

	Variable Sets			*Variable Sets*	
Model	*Inputs*	*Outputs*	*Model*	*Inputs*	*Outputs*
611	TOTAL_INC	TEACHING	621	TOTAL_EXP	TEACHING
		RESEARCH			RESEARCH
612	TOTAL_INC	TEACH_UGS	622	TOTAL_EXP	TEACH_UGS
		TEACH_PGS			TEACH_PGS
		RESEARCH			RESEARCH
613	HEFCS_INC	TEACHING	623	ACAD_DEPTS_EXP	TEACHING
	NON_HEFCS_INC	RESEARCH		CENTRAL_EXP	RESEARCH
614	HEFCS_INC	TEACH_UGS	624	ACAD_DEPTS_EXP	TEACH_UGS
	NON_HEFCS_INC	TEACH_PGS		CENTRAL_EXP	TEACH_PGS
		RESEARCH			RESEARCH

The variables used for Model 6 are defined as follows:

TOTAL_INC: Total income by institution in £ thousands, directly available from HESA.

HEFCS_INC: As above.

NON_HEFCS_INC: As above.

TOTAL_EXP: *Total expenditure* by institution in £ thousands, directly available from HESA.

ACAD_DEPTS_EXP: *Academic Departments* expenditure available directly from HESA statistics, it includes expenditure directly incurred by or on behalf of academic departments, which is not reimbursable by Research Councils or other bodies in respect of work carried out on their behalf. For 1995/96 there are 35 departmental cost centres to which this expenditure can be attributed.

CENTRAL_EXP: It is the sum of *Academic Services*, *Administration and Central Services*, *Premises*, *Residences and Catering Operations*, and *Research Grants and Contracts* expenditures, in £ thousands. *Academic Services* includes expenditure incurred on centralised academic services such as the Library, Learning Resource Centres, Central Computers, etc. *Administration and Central Services* includes expenditure incurred on central administration, general educational expenditure and staff and student facilities and amenities. *Premises* includes all expenditure incurred on the maintenance of premises and on roads and ground. *Residences and Catering* includes expenditure incurred in providing the residence, catering and any conference operations, including the cost of maintenance of residential and catering premises, salaries and any other identifiable costs relating to these operations. *Research Grants and Contracts* includes the total of the direct costs attributed to research grants and contracts.

The variables on the output side are as in Model 5.

In *Model 8: Efficiency Taking Account of Subject Mix*, to disentangle the subject-mix effect from the measurement of institutional efficiency, the four cost band groups proposed by HEFCE are considered: Group A: Clinical, Group B: Science, Engineering and Technology, Group C: Other high cost subjects with a studio, laboratory or fieldwork element, and Group D: All other subjects. The mapping of this groups to the information on expenditure by ACC, on FTE students by HESA subject categories, and FTE staff by UOA is given in Table 4-5.

Table 4-5: Subject Groups Based on Four Cost Bands

Cost bands	*Academic expenditure by 35 ACC*	*FTE students by HESA 19 subject areas*	*FTE staff by 69 UOA*
Group A: Clinical	01 Clinical Medicine, 02 Clinical Dentistry, 03 Veterinary Science	Medicine and dentistry, Veterinary science	01 Clinical Laboratory Sciences, 02 Community Based Clinical Subjects, 03 Hospital-Based Clinical Subjects
Group B: Science Engineering and Technology	04 Anatomy & Physiology, 08 Pharmacy, 09 Pharmacology, 10 Biosciences, 11 Chemistry, 12 Physics, 13 Agriculture & Forestry, 14 Earth Marine & Environmental Sciences, 15 General Sciences, 16 General Engineering, 17 Chemical Engineering, 18 Mineral Metallurgy & Materials Engineering, 19 Civil Engineering, 20 Electrical Electronic & Computer Engineering, 21 Mechanical Aero & Production Engineering, 22 Other Technologies	Biological sciences, Agriculture and related subjects, Physical sciences, Engineering and technology	04 Clinical Dentistry, 05 Pre-Clinical Studies, 06 Anatomy, 07 Physiology, 08 Pharmacology, 09 Pharmacy, 11 Other Studies and Professions Allied to Medicine, 12 Biochemistry, 14 Biological Sciences, 15 Agriculture, 16 Food Science and Technology, 17 Veterinary Science, 18 Chemistry, 19 Physics, 20 Earth Sciences, 21 Environmental Sciences, 26 General Engineering, 27 Chemical Engineering, 28 Civil Engineering, 29 Electrical and Electronic Engineering, 30 Mechanical Aeronautical and Manufacturing Engineering, 31 Mineral and Mining Engineering, 32 Metallurgy and Materials
Group C: Other high cost subjects with a studio laboratory or fieldwork element	05 Nursing & Paramedical Studies, 06 Health & Community Studies, 23 Architecture Built Environment & Planning, 24 Mathematics, 25 Information Technology & Systems Sciences, 26 Catering & Hospitality Management, 28 Geography, 33 Design & Creative Arts	Subjects allied to medicine, Mathematical sciences, Computer sciences, Architecture building and planning, Creative arts and design	10 Nursing, 13 Psychology, 22 Pure Mathematics, 23 Applied Mathematics, 24 Statistics and Operational Research, 25 Computer Science, 33 Built Environment, 34 Town and Country Planning, 35 Geography, 43 Business and Management Studies, 64 Art and Design, 65 Communication Cultural and Media Studies, 66 Drama Dance and Performing Arts, 67 Music, 68 Education, 69 Sports Related Subjects

Group D: All other subjects	07 Psychology & Behavioural Sciences, 27 Business & Management Studies, 29 Social Studies, 30 Librarianship Communication & Media Studies, 31 Language Based Studies, 32 Humanities, 34 Education, 41 Continuing Education	Social economic and political studies, Law, Business and administrative studies, Librarianship and information science, Languages, Humanities, Education, Combined studies,	36 Law, 37 Anthropology, 38 Economics and Econometrics, 39 Politics and International Studies, 40 Social Policy and Administration, 41 Social Work, 42 Sociology, 44 Accountancy, 45 American Studies (Canada the Caribbean Latin America and the USA), 46 Middle Eastern and African Studies, 47 Asian Studies, 48 European Studies, 49 Celtic Studies, 50 English Language and Literature, 51 French, 52 German Dutch and Scandinavian Languages, 53 Italian, 54 Russian Slavonic and East European Languages, 55 Iberian and Latin American Languages, 56 Linguistics, 57 Classics Ancient History Byzantine and Modern Greek Studies, 58 Archaeology, 59 History, 60 History of Art Architecture and Design, 61 Library and Information Management, 62 Philosophy, 63 Theology Divinity and Religious Studies

Model 8 was sub-divided into two models: Model 81 for universities with clinical studies and Model 82 for universities without clinical studies. Two models are necessary because of the need to have well-conditioned data matrices with the adopted weights restrictions. Unfortunately, this means that comparison of performance of universities with and without clinical studies is not possible.

Model 81: Universities with Clinical Studies

Variable Set			
Inputs	*Outputs*	*Weights Restrictions*	*Type*
GROUP_A_EXP	TEACH_A	TEACH_A + RES_A > GROUP_A_EXP	Virtual
GROUP_B_EXP	TEACH_B	TEACH_B + RES_B > GROUP_B_EXP	Virtual
GROUP_C_EXP	TEACH_C	TEACH_C + RES_C > GROUP_C_EXP	Virtual
GROUP_D_EXP	TEACH_D	TEACH_D + RES_D > GROUP_D_EXP	Virtual
CENTRAL_EXP	RES_A	GROUP_A_EXP > 4.5 GROUP_D_EXP	Absolute
	RES_B	GROUP_B_EXP > 2.0 GROUP_D_EXP	Absolute
	RES_C	GROUP_C_EXP > 1.5 GROUP_D_EXP	Absolute
	RES_D	GROUP_D_EXP > 0.05	Virtual
		CENTRAL_EXP > 0.40	Virtual

Model 82: Universities without Clinical Studies

Variable Set			
Inputs	*Outputs*	*Weights Restrictions*	*Type*
GROUP_B_EXP	TEACH_B	TEACH_B + RES_B > GROUP_B_EXP	Virtual
GROUP_C_EXP	TEACH_C	TEACH_C + RES_C > GROUP_C_EXP	Virtual
GROUP_D_EXP	TEACH_D	TEACH_D + RES_D > GROUP_D_EXP	Virtual
CENTRAL_EXP	RES_B	GROUP_B_EXP > 2.0 GROUP_D_EXP	Absolute
	RES_C	GROUP_C_EXP > 1.5 GROUP_D_EXP	Absolute
	RES_D	GROUP_D_EXP > 0.05	Virtual
		CENTRAL_EXP > 0.40	Virtual

The variables used in Model 8 are defined as follows:

GROUP_A_EXP: The sum of *Academic Departments* expenditure for ACCs belonging to Group A in £ thousands, from HESA statistics.

GROUP_B_EXP, GROUP_C_EXP, GROUP_D_EXP: As in GROUP_A_EXP, but for ACCs belonging to groups B, C, and D, respectively.

CENTRAL_EXP: It is the sum of *Academic Services, Administration and Central Services, Premises, Residences and Catering Operations*, and *Research Grants and Contracts* expenditures, in £ thousands, from HESA statistics.

TEACH_A: It is the total number of FTE students in the subject areas (from HESA statistics) belonging to Group A, as the measure of volume, multiplied by the average teaching rating for the university, as a measure of quality. In principle, the teaching output should be calculated for every UAQE, and the sum for all UAQE belonging to Group A obtained. However, as not all units have been assessed and FTE students are not available by UAQE, the average for each university is used instead.

TEACH_B, TEACH_C, TEACH_D: As in TEACH_A, but for subject areas belonging to groups B, C, and D, respectively.

RES_A: It is the total number of FTE academic staff in UOAs (from the RAE96 database) belonging to Group A, as the measure of volume, multiplied by the average research rating for the university, as measure of quality. In principle, the research output should be calculated for every UOA, and the sum for all UOA belonging to Group A, for each university, obtained. However, this information is not readily available.

RES_B, RES_C, RES_D: As in RES_A, but for UOA belonging to groups B, C, and D.

The weights restrictions reflect the following:

- Teaching and research output virtual weights combined for each group should be at least the same as the departmental expenditure virtual weight for that group. This reflects the fact that both departmental expenditure and a proportion of central expenditure contribute for the departmental output.

- The absolute weight for Group A should be at least four times that for Group D, for Group C 2.5 times, and for Group B 1.5 times. This reflects empirical evidence collected by HEFCE.
- The two last virtual weights restrictions reflect the evidence in the data; that for all units under assessment, at least 5% of the total expenditure is academic departmental expenditure in Group D, and at least 40% of the total expenditure is in central activities.

4.5.3 Results

The models in the previous section were solved using an output expansion orientation, which translates the funding agencies objective to calculate, given the resources each institution possesses, the level of activity they can deliver, both in terms of volume and quality. The results are in Appendix 4-2, where *COUNT* gives the number of units under consideration, *#EFF* is the number of units that appear efficient, *RADIAL* is the efficiency score, and the virtual weights for each factor are denoted by the name of the factor with the suffix *V*.

For Model 5 (see Table 4-9 and Table 4-10) concerned with the overall performance of the university from the state perspective, i.e. good quality teaching and research at value for money, plus evidence of autonomy from state appropriations: on average, English universities perform better than peripheral universities, as do old universities compared to new universities. These results are robust for different assumptions of returns to scale and different degrees of disaggregation of outputs. As to the difference in virtual weights for different outputs, English universities put more weight on attracting non-funding-council funds, than their non-English counterparts, the same for old compared with new universities. Both Oxford and Cambridge are efficient for all variants of Model 5. The most marked difference between old and new universities, is the weight given to research activity, this being almost negligible for new universities in all models. Old universities keep closer to the old rule of the UFC of a 1:2 ratio for research and teaching activity levels, respectively.

As for the efficiency given by Model 6, the results using the income variables, family Model 61x, on the input side (see Table 4-11 to Table 4-14), and using the expenditure variables, family Model 62x, on the input side (see Table 4-15 and Table 4-18) are not fully congruent. English universities, on average, are more efficient than peripheral universities for both families of models, under different assumptions of returns to scale, and different degrees of disaggregation of variables. However, as to the old/ new universities divide, the results are not conclusive. If the family Model 61x is considered, then old universities show, on average, to be more efficient than new universities. Whereas if the family Model 62x is used, then the results are not consistent for different returns to scale assumptions, and different degrees of disaggregation. Cambridge University is considered efficient for all variants of Model 6, whereas Oxford only achieves 100% efficiency for Model 623 and Model 624. As to the virtual weights assigned to the different outputs, again there is a marked difference between old and new universities for all variants of Model 6. Old universities have a more balanced distribution of virtual weights among their activities, whereas new universities give clear preference to teaching over research activity. If teaching output is disaggregated, it shows that the substantive difference is in the undergraduate teaching activity, as the weights given to postgraduate teaching are similar to those of the old universities. Strictly speaking, because of the possibility of alternative optimal weights in DEA, the above results cannot be interpreted accurately. However, the use of weights restrictions decreases the possibility of alternative weights. Also, we are comparing the averages of the optimal weights of all the universities rather than individual weights.

The results for Model 8, taking into account the effects of a university's subject mix in the measurement of efficiency, are shown in Table 4-19 for universities with clinical studies (Model 81), and in Table 4-20 for universities without clinical studies (Model 82). All 21 universities with clinical studies are old universities, 16 of which are English and 5 Non-English. The small number of DMUs, and the large number of input and output variables leads to lack of discrimination in the results, that even the imposed weights restrictions are not able to counteract: 15 units show 100% efficiency under CRS, and 20 under VRS, there is no visible difference between the efficiency and virtual weight allocation of English and Non-English universities. Oxford is efficient for both returns to scale assumption, whereas Cambridge is inefficient (96.98%) under CRS. The sample of universities without clinical studies is bigger, which improves the degree of discrimination of the results. When subject mix is taken into account, the difference in efficiency between English and Non-English universities decreases, and the new universities show, on average, to be more efficient than their old counterparts. The allocation of virtual weights between teaching and research activities remains, however, similar to previous models. As before, old universities, in general, give more weight to their research output and less to their teaching output than new universities.

4.5.4 Conclusion

Handling subject mix in the context of comparing universities' performance has been a problem. The use of weights restrictions that link the input-output divide to deal with subject-mix effects in this context, is a novel application of the method presented in Thanassoulis et al[33].

English universities seem to perform better, both when the State/ funding councils overall objectives are taken into consideration, and when efficiency is being studied, than more peripheral universities; confirming the conclusions of previous studies. English universities, as well as old universities seem to fare better in their autonomy, as they are able to attract more non-funding-council funds than non-English and new universities. Old universities, as expected from years of UGC funding[37] based on that teaching and research should be pursued in a ratio of approximately 2:1, have a more balanced portfolio of activities, whereas new universities concentrate heavily on the teaching activity. When disaggregation of outputs is considered, however, the weight given to postgraduate study does not seem to differ between the old and new sector, raising the question to what extent postgraduate teaching is, in fact, facilitated by departments pursuing research activity, or merely a natural extension of basic undergraduate study, that does not necessarily necessitate research activity to flourish. As to the comparison of efficiency between the old and new sector, the results are not conclusive when the subject mix factor is not taken into account, although the allocation of weights given to teaching and research remain similar. When subject mix is taken into account, a division between universities with clinical studies and without is necessary. The group of universities with clinical studies is relatively small, which leads to lack of discrimination in the results. Even so, Cambridge is one of the few universities that does show to be inefficient under CRS assumption. However, Oxford, with a similar funding structure shows up efficient. This incongruence in the results for Oxbridge, keeps the debate open as to the legitimacy of their different funding structure. As to the bigger group of universities without clinical studies, the difference in efficiency between English and Non-English universities decreases, which may lead to the conclusion that there is not after all a strong English/ periphery divide, but simply different subject-mix structures in their provision. New universities, also show up more efficient than old universities, when their subject mix structure is taken into account, even with the similar distribution of weights between teaching and research, as in models that do not take subject mix into account.

4.6 Concluding Remarks

The relationship between universities and the State is characterised by both increasing *accountability* to state appropriations for teaching and research, and the State's desire of increasing universities' *autonomy* from state funds. Funding councils, responsible for implementing governmental policy, undertake systematic evaluation of teaching and research activities to inform their allocation of resources to universities, and to inform the public at large.

The recent emphasis on a value-added approach to the evaluation of universities, makes DEA, with its ability of dealing with multiple inputs and outputs, a natural technique to be used in this context. It could inform peer judgement in the teaching quality and research assessment exercises. To this end, DEA models of the TQA and RAE are proposed. Unfortunately the lack of publicly available data has made it impossible to empirically test the models, and further the debate on the use of quantitative methods to inform qualitative evaluations. Moreover DEA could inform the allocation of resources on a value-for-money approach. This approach would be 'fairer' than the current consideration of unit cost in the allocation of new student places, or the output-oriented funding scale for research. A theoretical framework using DEA to assess the quality of teaching and research, and how it can be linked to the allocation or resources was proposed, but unfortunately could not be tested.

As to the comparison of total (i. e. teaching and research) university performance DEA models were also suggested, for the measurement of overall performance from the state perspective, and for efficiency, and then taking also into account the subject mix of each institution. Handling subject-mix in the context of university performance measurement has been a problem. The use of weights restrictions that link inputs and outputs concerning the same subject cost band, is a novel application in this context of the method proposed by Thanassoulis et al[33]. Although it is possible to solve these latter models, there are currently some caveats in the use of data currently readily available. The accuracy of the results obtained would be improved if the ideal measures discussed were used instead of the ones currently possible. This is especially significant for the implementation of Model 8, where a necessarily not completely accurate mapping of different units needed to be drawn.

Taking into consideration the previous caution, it seems that the use of DEA gives enough flexibility for the evaluation of a post-1992 heterogeneous group of universities with different identities and purposes; which are designed to be different kinds of institutions performing different missions. The currently used evaluative-funding models are charged to be biased towards the pre-1992 model of university, and up to now the merged sector has not been able to agree on a common set of quantitative performance measures. In fact, in an efficiency model that takes into account the subject mix of each university, the new universities seem to be performing better than the old ones, when both teaching and research activities are taken into consideration. However, when the relative success of a university becomes increasingly dependent upon its ability to secure resources not only from the State but also the market, the old universities seem to be better prepared for the challenge, as do English universities compared to those in the periphery.

A theme that is not fully resolved by the proposed models is the legitimacy of Oxbridge in securing extra resources from the State relative to the rest of the university sector.

It is hoped that the models proposed when solved with better quality data (which has started now to become available, and will probably become even more so, in the context of a market of higher education, where fee-paying students will increasingly play the role of 'clients') will help to shed more light in the issues dealt in here.

Appendix 4-1: Data

Table 4-6: Data for Model 5 and Model 6

			Inputs						Outputs			
DMU	PERIPHERY	NEW	HEFCS_INC	NON_HEFCS_INC	TOTAL_INC	ACAD_DEPTS_EXP	CENTRAL_EXP	TOTAL_EXP	TEACH_UGS	TEACH_PGS	TEACHING	RESEARCH
Anglia	0	1	29292	35031	64324	29576	33733	63309	242190	20132	262322	324
Aston	0	0	17741	21808	39549	17326	25516	42842	75133	11031	86164	1170
Bath	0	0	27691	38865	66556	25170	41150	66320	94956	25404	120360	4130
Birmingham	0	0	68611	135213	203824	89030	111691	200721	263143	87272	350415	10347
Bournemouth	0	1	17226	20020	37246	20953	15987	36940	140494	8487	148981	198
Bradford	0	0	24477	36665	61142	29034	30844	59878	108966	19381	128347	2702
Brighton	0	1	31456	32576	64032	32893	28999	61892	191408	19881	211290	1043
Bristol	0	0	56655	98153	154808	69737	86053	155790	191463	51655	243118	9051
Brunel	0	0	32377	40739	73116	32252	36789	69041	161874	41666	203540	2365
Cambridge	0	0	86314	180868	267182	101354	152557	253911	239816	95860	335676	23459
Central_England_Birmingham	0	1	34285	37510	71795	37358	34346	71704	237803	35890	273693	573
Central_Lancashire	0	1	35935	28812	64747	28610	34171	62781	231722	17606	249327	415
City	0	0	19149	44447	63596	33530	27563	61093	85895	42660	128555	1591
Coventry	0	1	36359	36114	72473	31606	38068	69674	231303	19008	250311	709
De_Monfort	0	1	55757	51156	106913	52242	52885	105127	226102	38148	264250	1768
Derby	0	1	20274	22792	43066	18729	23809	42538	171025	12336	183361	260
Durham	0	0	34948	53071	88019	34830	51641	86471	154909	44404	199313	4540
East_Anglia	0	0	25897	41853	67750	29924	36945	66869	129361	32310	161671	3624
East_London	0	1	30604	24819	55423	27365	28364	55729	177931	28177	206108	637
Essex	0	0	17824	30019	47843	18835	28510	47345	85674	21837	107511	2471
Exeter	0	0	32528	41743	74271	32936	40338	73274	132441	32785	165227	3065
Greenwich	0	1	39946	45148	85094	35614	49341	84955	221022	36737	257759	657
Hertfordshire	0	1	37746	47123	84869	40310	43188	83498	230029	21920	251950	713
Huddersfield	0	1	28172	25798	53970	26460	28377	54837	139814	26018	165832	713
Hull	0	0	27295	42364	69659	31819	39463	71282	135430	32349	167778	2952
Keele	0	0	17043	35179	52222	23250	30084	53334	98352	34906	133258	2262
Kent	0	0	22047	40906	62953	24803	38621	63424	126677	30671	157348	2577
Kings_College	0	0	47989	89099	137088	56771	78261	135032	178980	43772	222752	6369
Kingston	0	1	36430	34384	70814	34114	30831	64945	184053	25755	209808	614
Lancaster	0	0	30562	47975	78537	37206	52067	89273	143937	40094	184031	4302
Leeds_Metropolitan	0	1	38657	29736	68393	32386	35979	68365	205989	22475	228465	432
Leeds	0	0	75059	140380	215439	94804	118987	213791	309727	70697	380424	11165
Leicester	0	0	35748	73271	109019	45166	65141	110307	141794	87857	229650	5554
Lincolnshire_and_Humberside	0	1	25171	20291	45462	19971	22142	42113	191180	13866	205046	230
Liverpool_John_Moores	0	1	44632	39164	83796	42185	44724	86909	242922	24491	267413	993
Liverpool	0	0	62732	92930	155662	63133	92901	156034	211703	42515	254218	7206
Loughborough	0	0	34296	56455	90751	36610	50168	86778	148852	30956	179807	4048
Luton	0	1	18495	23221	41716	21656	21859	43515	206831	8708	215539	311
Manchester_Metropolitan	0	1	64784	48388	113172	55753	55193	110946	385152	56531	441683	1388
Manchester	0	0	86863	158062	244925	100825	148743	249568	290028	82865	372893	12992
Middlesex	0	1	40495	47200	87695	31500	52554	84054	287617	40833	328450	917
Newcastle	0	0	58667	89113	147780	58803	91640	150443	190861	51405	242266	8393
North_London	0	1	31590	23601	55191	21724	29454	51178	179165	19195	198360	545
Northumbria_at_Newcastle	0	1	38672	49940	88612	50390	39228	89618	271503	34487	305990	779
Nottingham_Trent	0	1	47634	51308	98942	52814	43600	96414	304831	22451	327282	940
Nottingham	0	0	55569	118564	174133	77811	97226	175037	253694	55689	309384	8550
Oxford_Brookes	0	1	25149	37999	63148	30066	32014	62080	146533	29511	176044	981
Oxford	0	0	85106	187175	272281	87615	173686	261301	239408	92817	332225	22229
Plymouth	0	1	41849	45216	87065	48313	37557	85870	300048	23823	323871	1328
Portsmouth	0	1	36652	41578	78230	45887	32877	78764	210621	18879	229500	1378
Reading	0	0	37493	63127	100620	47571	50899	98470	143791	57873	201665	4584
Royal_Holloway	0	0	18712	27711	46423	18105	27337	45442	89389	13376	102765	2259
Salford	0	0	23125	43984	67109	26084	39643	65727	105592	21579	127171	1906
Sheffield_Hallam	0	1	51138	43782	94920	47106	48704	95810	284237	36221	320458	1158
Sheffield	0	0	60985	117551	178536	73386	110206	183592	284826	95691	380517	9113
South_Bank	0	1	41478	43431	84909	43290	41400	84690	242524	54908	297432	622
Southampton	0	0	51780	102984	154764	54548	94612	149160	186714	47538	234252	7828
Staffordshire	0	1	31023	30053	61076	33353	33169	66522	213611	16778	230389	635
Sunderland	0	1	30017	29109	59126	28495	31035	59530	211462	12194	223656	454

			Inputs						Outputs			
DMU	PERIPHERY	NEW	HEFCS_INC	NON_HEFCS_INC	TOTAL_INC	ACAD_DEPTS_EXP	CENTRAL_EXP	TOTAL_EXP	TEACH_UGS	TEACH_PGS	TEACHING	RESEARCH
Surrey	0	0	24500	64052	88552	39466	42703	82169	102662	35716	138379	3433
Sussex	0	0	31091	39774	70865	29211	41604	70815	125675	34772	160447	4261
Teesside	0	1	23306	25222	48528	24078	22036	46114	149375	12834	162209	316
Thames_Valley	0	1	25681	30336	56017	31408	22062	53470	216974	22999	239973	161
UCL	0	0	89808	187863	277671	88645	182005	270650	195926	81761	277687	16495
Warwick	0	0	41241	93090	134331	50368	79842	130210	172950	89231	262181	7797
West_of_England_Bristol	0	1	43208	45971	89179	46600	40403	87003	284046	41124	325170	917
Westminster	0	1	42663	32356	75019	40330	36755	77085	183246	41158	224404	859
Wolverhampton	0	1	39523	41866	81389	35402	39823	75225	308710	29441	338151	421
York	0	0	23795	49918	73713	26516	45965	72481	97936	31505	129441	4553
Aberystwyth	1	0	24391	28120	52511	19908	32538	52446	111277	21799	133076	1766
Bangor	1	0	21382	35129	56511	25069	31891	56960	87581	26056	113637	1698
Cardiff	1	0	50263	63353	113616	46399	63891	110290	223740	52419	276159	6046
Glamorgan	1	1	27868	24422	52290	21928	28074	50002	204562	11836	216397	261
Swansea	1	0	30935	39847	70782	34899	37062	71961	131943	31835	163778	3004
Aberdeen	1	0	37911	58990	96901	37730	62284	100014	156265	25736	182001	4657
Abertay_Dundee	1	1	12578	8072	20650	9878	9731	19609	63864	5256	69120	188
Dundee	1	0	34210	46922	81132	33616	45939	79555	118724	24135	142859	3470
Edinburgh	1	0	85908	136870	222778	82503	136829	219332	272661	51964	324624	12819
Glasgow_Caledonian	1	1	33025	20105	53130	27298	24354	51652	172245	27330	199574	561
Glasgow	1	0	82943	101975	184918	80442	106465	186907	280629	45204	325833	8350
Heriot_Watt	1	0	23068	40576	63644	19717	42614	62331	74983	27525	102507	2087
Napier	1	1	27939	18162	46101	26339	18700	45039	128526	15067	143593	343
Paisley	1	1	22040	13479	35519	17141	17894	35035	102150	9359	111509	174
Robert_Gordon	1	1	21936	25658	47594	19568	28026	47594	113291	11549	124840	496
St_Andrews	1	0	21030	33953	54983	22491	33536	56027	101302	13118	114419	3165
Stirling	1	0	20649	29760	50409	19809	30187	49996	96647	18728	115375	1815
Strathclyde	1	0	59901	75681	135582	60154	74853	135007	216118	77334	293451	5099
Queens_Belfast	1	0	60035	61534	121569	51528	68340	119868	224083	54722	278805	4094
Ulster	1	0	55507	43038	98545	46186	59959	106145	256362	51371	307733	2835

Table 4-7: Data for Model 81

			Inputs					Outputs							
DMU	PERIPHERY	NEW	GROUP_A_EXP	GROUP_B_EXP	GROUP_C_EXP	GROUP_D_EXP	CENTRAL_EXP	TEACH_A	TEACH_B	TEACH_C	TEACH_D	RES_A	RES_B	RES_C	RES_D
Birmingham	0	0	17776	31689	7958	31607	111691	30714	119222	35038	207771	101	215	146	134
Bristol	0	0	20714	22185	6270	20568	86053	33487	71104	28264	155764	85	178	83	143
Cambridge	0	0	18585	41346	10696	30727	152557	37256	68256	39709	220954	78	277	142	232
Hull	0	0	991	6966	6566	17296	39463	111	35517	18090	141122	16	55	82	103
Keele	0	0	696	4973	5503	12078	30084	144	5923	20669	127171	9	46	56	92
Kings_College	0	0	11891	17943	12094	14843	78261	21214	39524	86793	91046	70	141	77	102
Leeds	0	0	19060	34488	11315	29941	118987	31240	131200	38940	210750	88	201	165	135
Leicester	0	0	10586	12692	2873	19015	65141	17157	47362	6940	174082	74	87	58	94
Liverpool	0	0	18453	21710	9994	12976	92901	37989	73610	43462	129947	81	195	99	98
Manchester	0	0	24658	29259	11729	35179	148743	41759	83314	59986	217025	132	211	181	200
Newcastle_upon_Tyne	0	0	14953	22335	8520	12995	91640	25970	83163	39406	116698	115	146	106	93
Nottingham	0	0	12312	28461	17660	19378	97226	22695	97961	76133	168943	89	170	126	107
Oxford	0	0	13617	34607	6440	32951	173686	17819	87291	30910	238564	129	238	110	327
Sheffield	0	0	17187	20457	14470	21272	110206	32886	108178	88146	197745	100	168	134	138
Southampton	0	0	7165	23265	10182	13936	94612	17425	86459	48155	113611	52	178	119	107
UCL	0	0	33289	24651	11435	19270	182005	36144	93283	61348	104667	143	200	117	162
Aberdeen	1	0	8019	15513	3472	10726	62284	18470	68860	24280	94880	65	85	51	92
Dundee	1	0	10080	8150	8210	7176	45939	19778	27121	42835	63584	60	88	93	56
Edinburgh	1	0	23379	24585	9527	25012	136829	36920	95118	31793	181293	114	238	118	165
Glasgow	1	0	25350	27325	7867	19900	106465	47472	92152	32693	174103	122	210	116	148
Queens_Belfast	1	0	11513	18469	6173	15373	68340	24070	78028	38180	177342	54	169	59	114

Table 4-8: Data for Model 82

			Inputs				*Outputs*					
UNITS	*PERIPHERY*	*NEW*	*GROUP_B_EXP*	*GROUP_C_EXP*	*GROUP_D_EXP*	*CENTRAL_EXP*	*TEACH_B*	*TEACH_C*	*TEACH_D*	*RES_B*	*RES_C*	*RES_D*
Anglia_Polytechnic	0	1	4238	14918	10421	33733	29683	84518	171917	41	155	87
Aston_Birmingham	0	0	9014	1140	7172	25516	18611	16521	59632	51	45	9
Bath	0	0	13490	3045	8245	41150	47889	28905	57451	102	65	39
Bournemouth	0	1	5642	10454	4857	15987	24606	57393	72027	27	58	34
Bradford	0	0	14614	3675	10745	30844	47196	23259	74400	79	56	47
Brighton	0	1	5976	15401	11516	28999	19670	84277	116445	63	203	39
Brunel	0	0	12230	8137	11885	36789	73516	36357	112957	67	67	33
Central_England_Birmingham	0	1	5527	18837	12994	34346	33476	119381	116559	56	236	82
Central_Lancashire	0	1	7470	9987	11153	34171	51173	65179	148946	55	119	82
City	0	0	7895	11819	13816	27563	21782	62702	82690	49	146	29
Coventry	0	1	11481	10876	9249	38068	103496	60684	92888	90	147	39
De_Monfort	0	1	12613	22235	17394	52885	57480	146007	147251	111	211	124
Derby	0	1	3532	6863	8334	23809	29463	40763	119289	56	90	48
Durham	0	0	10933	4315	19582	51641	51459	16546	144092	81	103	113
East_Anglia	0	0	9251	8512	12161	36945	32815	32495	102422	71	65	84
East_London	0	1	7625	6714	13026	28364	49538	58459	91932	49	124	73
Essex	0	0	5390	2154	11291	28510	23196	10850	93139	38	20	96
Exeter	0	0	9426	4592	18536	40338	34642	11347	143958	66	94	88
Greenwich	0	1	8353	14867	12394	49341	53390	86412	131157	80	161	59
Hertfordshire	0	1	13569	15610	11131	43188	73510	82122	105691	107	177	51
Huddersfield	0	1	5450	10802	10208	28377	37729	50609	110676	52	115	54
Kent	0	0	6708	4663	13432	38621	29693	15715	129782	47	68	115
Kingston	0	1	9494	12897	11723	30831	57825	61321	107113	57	165	51
Lancaster	0	0	11330	1944	23932	52067	33415	14781	161366	54	94	95
Leeds_Metropolitan	0	1	2405	15557	14424	35979	27000	79843	141487	56	165	98
Lincolnshire_and_Humberside	0	1	4812	6543	8616	22142	38546	51406	113787	33	97	47
Liverpool_John_Moores	0	1	11602	13964	16619	44724	60744	81369	148393	102	140	102
Loughborough	0	0	19332	5493	11785	50168	75832	18365	97696	105	112	48
Luton	0	1	3360	6968	11328	21859	24549	39217	140407	32	116	66
Manchester_Metropolitan	0	1	14229	16589	24935	55193	76016	80968	295605	112	319	170
Middlesex	0	1	3683	15501	12316	52554	34481	82862	212385	45	189	97
North_London	0	1	5197	5528	10999	29454	21188	30682	153217	31	66	73
Northumbria_at_Newcastle	0	1	6406	26233	17751	39228	46508	106733	158145	42	291	111
Nottingham_Trent	0	1	15017	16123	21674	43600	81607	96183	195667	66	204	100
Oxford_Brookes	0	1	7666	12111	10289	32014	26529	63749	100764	50	114	38
Plymouth	0	1	16070	19020	13215	37557	116983	83971	125277	100	132	72
Portsmouth	0	1	15314	15267	15306	32877	64751	75184	107305	92	183	84
Reading	0	0	20629	8397	18545	50899	58771	25313	144766	93	138	93
Royal_Holloway	0	0	5385	4891	7829	27337	27423	18167	64301	39	64	63
Salford	0	0	11342	7775	6967	39643	59050	28074	59914	82	43	45
Sheffield_Hallam	0	1	13735	18386	14985	48704	55377	97363	174891	95	286	62
South_Bank	0	1	11075	19422	12793	41400	63341	86844	134791	111	176	80
Staffordshire	0	1	6352	12973	14028	33169	45717	68011	123582	65	132	94
Sunderland	0	1	9866	6762	11867	31035	57983	68472	98754	72	117	45
Surrey	0	0	19494	10563	9409	42703	60421	36947	76950	115	64	29
Sussex	0	0	10570	3743	14757	41604	52448	18818	116784	91	70	126
Teesside	0	1	5314	11217	7547	22036	28722	59585	91555	47	145	46
Thames_Valley	0	1	789	16722	13897	22062	3001	90863	169244	6	30	33
Warwick	0	0	17687	3572	28882	79842	78495	32132	209356	92	146	133
West_of_England_Bristol	0	1	8534	19378	18688	40403	65000	94796	156763	83	257	103
Westminster	0	1	5979	13186	21165	36755	36417	86608	138935	60	163	78
Wolverhampton	0	1	5663	14082	15657	39823	33950	100933	225634	38	232	83
York	0	0	9217	6026	11273	45965	39006	16861	81852	65	53	103
Aberystwyth	1	0	6682	3886	9340	32538	39362	16685	83048	48	54	80
Bangor	1	0	9591	4611	10867	31891	44996	16345	71181	79	82	45
Cardiff	1	0	18661	7703	20035	63891	68182	47832	181181	137	130	104
Glamorgan	1	1	6240	7158	8387	28074	43517	52906	136550	48	97	70
Swansea	1	0	10647	8311	15941	37062	52445	26803	112544	68	103	86
Abertay_Dundee	1	1	4523	2050	3305	9731	21188	17550	33210	33	44	25
Glasgow_Caledonian	1	1	7992	9584	9722	24354	34922	63137	102648	68	125	48
Heriot_Watt	1	0	12904	1795	5018	42614	46893	24196	60838	86	57	18
Napier	1	1	9615	7355	9369	18700	44925	31469	70450	67	107	44

	Inputs						Outputs					
UNITS	PERIPHERY	NEW	GROUP_B_EXP	GROUP_C_EXP	GROUP_D_EXP	CENTRAL_EXP	TEACH_B	TEACH_C	TEACH_D	RES_B	RES_C	RES_D
Paisley	1	1	7230	5097	4814	17894	26740	11011	91050	48	78	32
Robert_Gordon	1	1	7497	7950	4121	28026	30330	51656	50209	67	100	34
St_Andrews	1	0	9528	1971	10992	33536	28728	8577	73966	54	39	81
Stirling	1	0	3001	1477	14225	30187	23837	6630	102877	34	64	72
Strathclyde	1	0	23970	7779	28405	74853	86891	41540	221048	129	199	85
Ulster	1	0	11174	15076	19936	59959	55986	98709	157713	80	248	79

Appendix 4-2: Results

Table 4-9: Results for Model 51

		COUNT	#EFF	RADIAL	HEFCS_INC_V	NON_HEFCS_INC_V	TEACHING_V	RESEARCH_V
CRS	All	89	6	72.95	141.87	32.57	53.34	14.09
	English	69	6	75.75	136.00	34.60	52.06	13.34
	Non-English	20	0	63.27	162.13	25.55	57.76	16.69
	Old	49	5	78.29	130.94	32.73	42.67	24.60
	New	40	1	66.40	155.27	32.37	66.41	1.22
VRS	All	89	16	82.26	102.66	23.93	64.95	11.12
	English	69	15	84.96	96.06	25.32	63.60	11.08
	Non-English	20	1	72.95	125.40	19.14	69.59	11.27
	Old	49	12	83.55	108.12	32.25	48.06	19.69
	New	40	4	80.69	95.96	13.74	85.64	0.62

Table 4-10: Results for Model 52

		COUNT	#EFF	RADIAL	HEFCS_INC_V	NON_HEFCS_INC_V	TEACH_UGS_V	TEACH_PGS_V	RESEARCH_V
CRS	All	89	8	75.18	137.07	25.49	47.93	13.50	13.08
	English	69	8	78.02	131.52	26.30	47.37	13.57	12.76
	Non-English	20	0	65.40	156.24	22.67	49.85	13.26	14.21
	Old	49	7	79.97	128.02	33.44	34.45	8.43	23.68
	New	40	1	69.32	148.16	15.75	64.44	19.70	0.10
VRS	All	89	18	83.86	97.47	22.12	56.32	11.28	10.28
	English	69	17	86.42	92.53	23.44	55.91	10.18	10.47
	Non-English	20	1	75.01	114.55	17.57	57.75	15.04	9.64
	Old	49	12	85.00	100.02	30.71	39.67	11.29	18.33
	New	40	5	82.45	94.36	11.60	76.72	11.26	0.43

Table 4-11: Results for Model 611

		COUNT	#EFF.	RADIAL	TOTAL_INC_V	TEACHING_V	RESEARCH_V
CRS	All	89	2	75.08	135.68	66.25	33.75
	English	69	2	76.25	133.18	66.62	33.38
	Non-English	20	0	71.03	144.29	65.01	34.99
	Old	49	1	77.24	132.14	47.46	52.54
	New	40	1	72.42	140.01	89.28	10.72
VRS	All	89	4	81.97	106.31	72.87	27.13
	English	69	3	83.31	101.55	73.52	26.48
	Non-English	20	1	77.34	122.76	70.63	29.37
	Old	49	1	82.41	110.34	55.52	44.48
	New	40	3	81.43	101.37	94.12	5.88

Table 4-12: Results for Model 612

		COUNT	#EFF.	RADIAL	TOTAL_INC_V	TEACH_UGS_V	TEACH_PGS_V	RESEARCH_V
CRS	All	89	5	83.17	122.06	46.25	28.78	24.97
	English	69	5	84.62	119.81	46.51	29.18	24.31
	Non-English	20	0	78.17	129.81	45.36	27.40	27.24
	Old	49	3	83.42	121.68	31.87	25.81	42.32
	New	40	2	82.87	122.53	63.87	32.42	3.71
VRS	All	89	13	86.99	106.38	52.14	26.56	21.31
	English	69	11	88.07	102.73	52.98	25.96	21.07
	Non-English	20	2	83.26	118.97	49.24	28.63	22.13
	Old	49	7	87.31	102.41	36.19	28.33	35.48
	New	40	6	86.60	111.24	71.67	24.39	3.94

Table 4-13: Results for Model 613

		COUNT	#EFF.	RADIAL	HEFCS_INC_V	NON_HEFCS_INC_V	TEACHING_V	RESEARCH_V
CRS	All	89	6	81.41	44.59	80.15	66.23	33.77
	English	69	4	81.69	47.63	76.46	67.62	32.38
	Non-English	20	2	80.44	34.11	92.90	61.42	38.58
	Old	49	3	82.55	60.96	61.96	48.49	51.51
	New	40	3	80.01	24.53	102.43	87.96	12.04
VRS	All	89	14	86.04	50.01	55.82	72.03	27.97
	English	69	11	86.79	54.06	48.38	73.55	26.45
	Non-English	20	3	83.45	36.06	81.51	66.82	33.19
	Old	49	8	86.36	53.28	53.84	55.42	44.58
	New	40	6	85.66	46.01	58.25	92.38	7.62

Table 4-14: Results for Model 614

		COUNT	#EFF.	RADIAL	HEFCS_INC_V	NON_HEFCS_INC_V	TEACH_UGS_V	TEACH_PGS_V	RESEARCH_V
CRS	All	89	11	85.63	44.07	74.34	48.94	22.53	28.54
	English	69	2	86.32	49.87	67.49	49.20	24.11	26.69
	Non-English	20	9	83.25	24.03	97.97	48.02	17.06	34.92
	Old	49	6	85.65	48.88	69.46	33.29	21.14	45.57
	New	40	5	85.61	38.17	80.31	68.10	24.22	7.67
VRS	All	89	19	89.19	43.35	58.24	54.49	21.97	23.54
	English	69	15	90.01	49.06	50.65	55.13	22.81	22.06
	Non-English	20	4	86.35	23.66	84.45	52.31	19.04	28.64
	Old	49	11	89.30	44.66	55.67	37.54	24.21	38.26
	New	40	8	89.05	41.75	61.40	75.27	19.22	5.52

Table 4-15: Results for Model 621

		COUNT	#EFF.	RADIAL	TOTAL_EXP_V	TEACHING_V	RESEARCH_V
CRS	All	89	2	76.25	133.44	68.51	31.49
	English	69	2	77.56	130.79	68.82	31.18
	Non-English	20	0	71.74	142.59	67.47	32.53
	Old	49	1	76.17	133.75	50.61	49.39
	New	40	1	76.36	133.07	90.45	9.55
VRS	All	89	5	81.35	110.28	74.13	25.87
	English	69	4	82.70	105.04	74.59	25.41
	Non-English	20	1	76.66	128.37	72.55	27.46
	Old	49	1	80.71	108.46	59.06	40.94
	New	40	4	82.12	112.52	92.59	7.41

Table 4-16: Results for Model 622

		COUNT	#EFF.	RADIAL	TOTAL_EXP_V	TEACH_UGS_V	TEACH_PGS_V	RESEARCH_V
CRS	All	89	8	83.09	122.26	47.56	28.37	24.06
	English	69	8	84.57	119.96	47.70	28.84	23.46
	Non-English	20	0	77.96	130.21	47.09	26.76	26.14
	Old	49	3	82.47	123.14	33.20	27.59	39.21
	New	40	5	83.84	121.19	65.16	29.34	5.50
VRS	All	89	14	86.52	107.70	51.77	27.54	20.69
	English	69	12	87.68	104.06	52.53	27.18	20.28
	Non-English	20	2	82.51	120.26	49.13	28.76	22.11
	Old	49	6	86.69	102.86	36.21	30.69	33.10
	New	40	8	86.31	113.63	70.82	23.68	5.50

Table 4-17: Results for Model 623

		COUNT	#EFF.	RADIAL	ACAD_DEPTS_EXP_V	CENTRAL_EXP_V	TEACHING_V	RESEARCH_V
CRS	All	89	6	80.39	61.14	65.09	68.48	31.52
	English	69	6	81.27	55.90	68.96	68.59	31.41
	Non-English	20	0	77.35	79.21	51.73	68.09	31.91
	Old	49	2	80.00	67.37	59.08	51.52	48.48
	New	40	4	80.86	53.51	72.44	89.25	10.75
VRS	All	89	10	85.18	56.10	54.46	72.75	27.25
	English	69	9	85.81	48.56	58.44	73.03	26.97
	Non-English	20	1	83.01	82.11	40.72	71.79	28.21
	Old	49	2	84.81	65.23	45.12	57.45	42.55
	New	40	8	85.63	44.92	65.89	91.50	8.50

Table 4-18: Results for Model 614

		COUNT	#EFF.	RADIAL	ACAD_DEPTS _EXP_V	CENTRAL_EX P_V	TEACH_UGS_ V	TEACH_PGS_ V	RESEARCH_V
CRS	All	89	10	85.62	56.14	62.31	49.28	25.29	25.43
	English	69	10	86.76	53.52	63.32	49.43	25.75	24.83
	Non-English	20	0	81.71	65.16	58.83	48.79	23.70	27.51
	Old	49	5	85.25	56.53	62.35	33.15	26.30	40.55
	New	40	5	86.09	55.66	62.25	69.04	24.04	6.91
VRS	All	89	23	89.36	51.42	54.59	53.07	24.88	22.04
	English	69	20	89.89	46.03	56.95	53.77	24.56	22.06
	Non-English	20	3	87.56	70.03	46.47	50.66	26.01	23.33
	Old	49	12	89.75	54.39	47.58	37.01	28.37	34.62
	New	40	11	88.90	47.80	63.18	72.75	20.61	6.64

Table 4-19: Results for Model 81

		COUNT	#EFF.	RADIAL	GROUP_A_EXP_V	GROUP_B_EXP_V	GROUP_C_EXP_V	GROUP_D_EXP_V	CENTRAL_EXP_V	TEACH_A_V	TEACH_B_V	TEACH_C_V	TEACH_D_V	RES_A_V	RES_B_V	RES_C_V	RES_D_V
CRS	All	21	15	98.28	22.99	15.92	15.33	5.62	42.17	12.37	9.50	10.26	2.30	18.79	20.69	12.81	13.29
	Engl.	16	11	98.14	21.45	15.41	16.74	5.82	42.84	12.41	10.21	10.47	2.39	17.24	18.45	14.61	14.23
	N. E.	5	4	98.74	27.92	17.58	10.84	5.00	40.00	12.25	7.24	9.61	2.00	23.74	27.85	7.04	10.27
VRS	All	21	20	99.95	20.95	12.58	10.97	5.60	40.49	11.26	9.62	10.40	7.07	20.92	14.94	11.54	14.25
	Engl.	16	15	99.93	20.05	12.45	12.66	5.78	40.64	11.63	9.84	12.11	6.68	19.05	14.25	11.70	14.76
	N. E.	5	5	100.00	23.81	12.99	5.55	5.00	40.00	10.07	8.94	4.95	8.34	26.90	17.15	11.03	12.62

Table 4-20: Results for Model 82

		COUNT	#EFF.	RADIAL	GROUP_B_EXP_V	GROUP_C_EXP_V	GROUP_D_EXP_V	CENTRAL_EXP_V	TEACH_B_V	TEACH_C_V	TEACH_D_V	RES_B_V	RES_C_V	RES_D_V
CRS	All	68	31	93.16	21.55	20.96	7.33	58.50	15.37	17.22	10.58	22.16	17.06	17.61
	Engl.	53	24	93.06	19.40	21.26	7.11	60.80	16.03	19.89	11.66	19.14	16.31	16.97
	NonE.	15	7	93.49	29.16	19.93	8.11	50.39	13.06	7.77	6.76	32.82	19.69	19.90
	Old	28	7	87.52	23.65	24.05	8.61	59.23	11.18	7.54	11.31	29.32	19.08	21.58
	New	40	24	97.11	20.09	18.80	6.44	57.99	18.31	24.00	10.07	17.14	15.64	14.84
VRS	All	68	51	97.37	15.00	18.87	6.32	51.58	12.26	18.59	14.13	27.37	10.50	17.15
	Engl.	53	39	97.49	14.33	18.27	6.26	51.74	11.53	20.11	15.90	26.47	9.46	16.53
	NonE.	15	12	96.94	17.37	21.00	6.52	51.01	14.88	13.21	7.85	30.53	14.17	19.36
	Old	28	17	94.94	17.93	20.02	7.43	53.77	11.29	8.17	9.06	33.22	14.73	23.53
	New	40	34	99.07	12.95	18.07	5.53	50.05	12.94	25.89	17.67	23.27	7.54	12.69

4.7 References

[1] Great Britain, Department of Education and Science (1991). *Higher Education: A New Framework (Cm1541).* HMSO: London.

[2] Higher Education Funding Council for England (1994). *Introduction to HEFCE.* HEFCE: Bristol.

[3] Salter B and Tapper T (1994). *The State and Higher Education.* The Woburn Press: Essex.

[4] Great Britain, Committee on Higher Education, Lord Robbins (chairman) (1963). *Higher Education, Report of the Committee on Higher Education Under the Chairmanship of Lord Robbins 1961-63 (Cmnd 2154).* HMSO: London.

[5] Great Britain (1992). *The Further and Higher Education Act 1992.* HMSO: London.

[6] Great Britain, National Committee of Inquiry into Higher Education, Sir Ron Dearing (1997). *Higher Education in the Learning Society, Report of the National Committee.* HMSO: Norwich.

[7] Great Britain, Department of Education and Science (1987). *Higher Education: Meeting the Challenge (Cm114).* HMSO: London.

[8] Cave M, Hanney S, Kogan M, and Trevett G (1988). *The Use of Performance Indicators in Higher Education: A Critical Analysis of Developing Practice.* Jessica Kingsley: London.

[9] Committee of Vice-Chancellors and Principals of the Universities of the United Kingdom, Steering Committee for Efficiency Studies in Universities, Sir Alex Jarratt (1985). *Report of the Steering Committee for Efficiency Studies in Universities.* CVCP: London.

[10] Committee of Vice-Chancellors and Principals of the Universities of the United Kingdom (1987). *University Management Statistics and Performance Indicators.* CVCP: London.

[11] Higher Education Funding Council for England (1994). *Units of Assessment of the Quality of Education, C14/94.* HEFCE: Bristol.

[12] Higher Education Funding Council for England (1995). *Quality Assessment Between October 1996 and September 1998, C26/95.* HEFCE: Bristol.

[13] Higher Education Funding Council for England (1994). *1996 Research Assessment Exercise: Units of Assessment, RAE96 3/94.* HEFCE: Bristol.

[14] Higher Education Funding Council for England (1994). *Conduct of the Research Assessment Exercise: Panel Membership and Units of Assessment, RAE96 2/94.* HEFCE: Bristol.

[15] Higher Education Funding Council for England (1995). *A Guide to Funding Higher Education in England, M13/95.* HEFCE: Bristol.

[16] Higher Education Funding Council for England (1996). *Funding Method for Research, CP2/96.* HEFCE: Bristol.

[17] Higher Education Funding Council for England (1996). *Funding Method for Teaching, CP1/96.* HEFCE: Bristol.

[18] Higher Education Funding Council for England (1996). *Funding Method for Teaching from 1998-99, C21/96*. HEFCE: Bristol.

[19] Higher Education Funding Council for England (1993). *Accountability for Research Funds, CP4/93*. HEFCE: Bristol.

[20] Higher Education Funding Council for England (1994). *Overview of Recent Developments in HE, M2/94*. HEFCE: Bristol.

[21] Higher Education Funding Council for England (1997). *Research Assessment: Consultation, RAE 2/97*. HEFCE: Bristol.

[22] Doyle JR, Arthurs AJ, Green RH, McAulay L, Pitt MR, Bottomley PA, and Evans W (1996). The judge, the model of the judge, and the model of the judged as judge: analyses of the UK 1992 research assessment exercise data for business management studies. *Omega* **24:** 13-28.

[23] Dyson RG, Allen R, Shale E and Santos AR (1997). Pitfalls and protocols in the use of DEA. Presented at the *39th Annual Conference of the Operational Research Society*. September 1997, Bath, UK.

[24] Johnes J and Johnes G (1995). Research funding and performance in UK university departments of economics: a frontier analysis. *Econ Educ* **14**: 301-314.

[25] Beasley JE (1990). Comparing university departments. *Omega* **18:** 171-183.

[26] Johnes G and Johnes J (1993). Measuring the research of UK economics departments: an application of data envelopment analysis. *Oxf Econ Pap* **45**: 332-347.

[27] Johnes G (1995). Scale and technical efficiency in the production of economic research. *Appl Econ Lett* **2**: 7-11.

[28] Committee of Vice-Chancellors and Principals of the Universities of the United Kingdom (1995). *Higher Education Management Statistics: A Future Strategy*. CVCP: London

[29] Higher Education Funding Council for England (1994). *Accountability for Research Funds, C4/94.* HEFCE: Bristol.

[30] Higher Education Funding Council for England (1997). *Research Activity Survey, 97/21.* HEFCE: Bristol.

[31] Johnes J and Taylor J (1990). *Performance Indicators in Higher Education.* Open University Press: Buckingham.

[32] Swain H (1997). Blackstone fudge on college fees. *The Times Higher Education Supplement* **Nov. 14**: 1 and 48.

[33] Thanassoulis E, Boussofiane A and Dyson RG (1995). Exploring output quality targets in the provision of perinatal care in England using data envelopment analysis. *Euro J Opl Res* **80**: 588-607.

[34] Higher Education Statistics Agency (1997). *Students in Higher Education Institutions 1995/96.* HESA: Cheltenham.

[35] Higher Education Statistics Agency (1997). *Resources in Higher Education Institutions 1995/96.* HESA: Cheltenham.

[36] O'Leary J (1997). *The Times Good University Guide 1997.* Times Books: London.

[37] Great Britain, Department of Education and Science (1987). *Review of the University Grants Committee (Cmnd. 81).* HMSO: London.

5. THE INSTITUTIONAL PERSPECTIVE

5.1 Introduction

The universities write their own mission statements, develop their own plans, and define their own strategies for achieving their goals. But, needless to say, they undertake these tasks in circumstances not of their own choosing. Moreover, they have to be sensitive to the fact that these circumstances may change, which could require them to modify their strategies. If diversity is the main characteristic of the new university system model, then the constant changing of its context is the main pressure to which it has to adjust.

In this thesis three different classes of stakeholders interested in universities performance measurement were differentiated. The wider society represented by Government, whose policy is implemented by the funding councils; the applicant who is in the process of choosing a university; and the institutions themselves, 'sandwiched' between the external assessment by the two former classes and their own objectives. The present chapter will focus on the institutional level, and the contribution of DEA to inform management is explored and illustrated in an application to the University of Warwick.

Dyson et al (1994)[1] show how definitions of efficiency in DEA may conflict with the mission of the organisation being assessed. Only by taking into account the mission and objectives of the organisation can DEA have an effectiveness orientation.

Different institutions will have different objectives. But, in general, all will want, first of all, to attract students; then to develop competencies in them, with different emphases on teaching and research, regional and national scope, vocational and comprehensive courses; and, finally, regardless of their mission, they will want to keep financially sound to survive as an organisation. As to the purposes of measuring performance, there is a summative purpose, for instance, when allocating resources, but also a formative one in the search for enhanced performance.

The evaluation that a university makes of its departments will be considered, taking into consideration the objectives of the university itself, and its external evaluation by applicants and the State. It is important to note that direct comparison of departments of different disciplines is not a legitimate exercise. Thus in this study the departments are initially compared to their counterparts in other universities, in an external benchmarking exercise; and only afterwards are the departments' performances compared with other departments at the university being studied.

In the following section a theoretical framework for institutional performance measurement is described, that takes into account the interfaces between its internal and external environments. A section describing the application of the methodology to the University of Warwick then follows, and, finally, some conclusions are derived.

The application to the University of Warwick was not commissioned by any body of the University, nor are the results necessarily condoned by its authorities. However, part of the data for the application was supplied by the University and authorisation was obtained from the Registrar to publish the results.

5.2 A Framework for Institutional Performance Management

The framework developed in this chapter supports the analysis of the performance by a university of its departments taking account of both external and internal dimensions. The idea of analysing the performance of an organisation as a portfolio of businesses (in the present case a portfolio of departments belonging to a university), dates back to a technique to support strategic option formulation, the Boston Consulting Group (BCG) matrix[2] developed in the 1960s (see Figure 5-1). The BCG matrix views the corporation as a portfolio of businesses, each one offering a unique contribution to an externally dependent dimension - market growth rate, and an internal performance dependent dimension - relative market share. The assumptions here being that the corporation's objectives are growth and profitability, and that these are determined by market share and market size. The advantage of addressing the different units within the corporation as having to some extent independent performances in relation to market growth and market share is the ability to transfer resources from worse performers to more promising ones. Hence increasing the aggregate performance of the firm. All the businesses of the firm are located in the four-quadrant grid, offering a compact picture of their strengths and weaknesses. It identifies the capacity of each business to generate cash and its requirements for cash. The *stars* are best performers, they have high market growth and high relative market share. They generate large amounts of cash, but also require significant amounts of cash to sustain their competitiveness. As a result the net cash flow is modest. The *cash cows*, on the other hand, generate large positive cash flow for the organisation, because of their very high competitive strength in static or declining markets. The *question marks* are attractive because of the high market growth, however, they still have not achieved significant shares of their markets. The net cash flow is likely to be significantly negative. Before deciding to invest or divest in these businesses, the firm must decide if they are worth promoting or liquidating, according to the nature of the competition. As to the *dogs*, the conclusion is inevitably to divest. If they do generate net cash flows, they are mainly to maintain their own operations, and are not adding up value to the firm. Ideally a business would evolve from a question mark to a star, and eventually to a cash cow, which can then generate the funds to invest in new attractive question marks that go on to be stars, in a continuous healthy life circle.

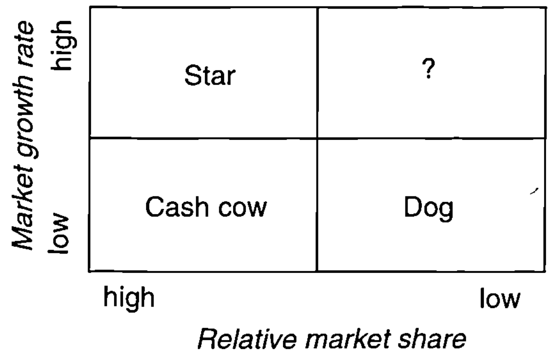

Figure 5-1: Growth - share matrix

Adopting and transforming the portfolio approach, Dyson et al. (1990)[3] and Boussofiane et al. (1991)[4] showed how DEA can be used to generate efficiency vs. profitability matrices (see Figure 5-2). In this case it is argued that, even for profit-making organisations, profitablility should not be the only measure of performance, the reason being the effect that environmental factors outside the control of the firm might have on its performance. In deciding about the performance of units, and how to allocate resources between them, both efficiency and profitability are relevant. A profitable unit may be managed efficiently or simply be enjoying favourable environmental factors, while an unprofitable unit may be badly managed or simply be experiencing adverse conditions. *Star* units are best performers in both dimensions, and are possibly operating under favourable conditions. The *sleepers*, however, although profitable, are inefficient. They have a potential for greater efficiency and thus for increased profits. Units located on the *question mark* quadrant are under-performing on both performance dimensions, having the potential for increased profitability as well as efficiency. Finally, *dog* units are efficient but have low profitability, presumably due to unfavourable environmental conditions. Thus, it might make sense to divest these units and reallocate the resources to more promising ones.

Athanassopoulos and Thanassoulis (1995)[5] have developed a similar concept with reference to the pubs of a major brewery. Market efficiency reflected the ability to attract customs and profitability measured as profit per pound of turnover. This analysis had the advantage of disentangling performance due to environmental factors and performance due to quality of management. Another advantage of the analysis is the possibility of benchmarking, where the practice of pubs showing good performance can be disseminated across the sector.

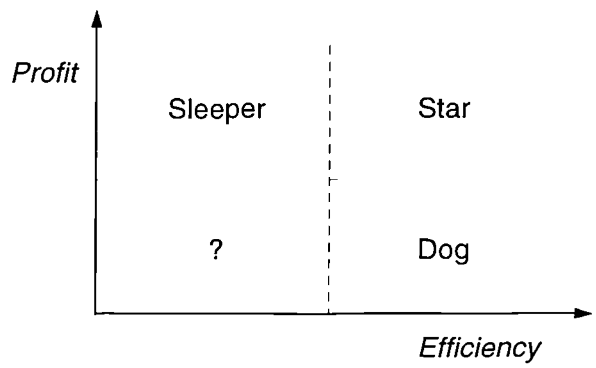

Figure 5-2: The efficiency - profitability matrix

In this chapter the use of matrices of different dimensions of performance is further explored for a portfolio of departments of a university. The methodology considers two interfaces between internal and external environments: the interface between the applicant and the institution, and the interface between the State and the institution. In the first case, a dimension of performance - *reality* - for the department, in what it has to offer to the applicant, is taken into account in conjunction with a second dimension of performance - *image* - of the department as perceived by the market - potential applicant. In the second case, a dimension of performance - *accountability* - representing the performance of the department in complying with objectives defined by the State, when allocating funds to the university, is analysed in association with a dimension of performance - *autonomy* - which represents the contribution of the department to the autonomy of the institution in the pursuit of its own objectives, regardless of the State's.

The measurement of the dimensions under consideration for a department is initially in relation to other departments of the same discipline in other universities. This approach represents a departure from other studies (see for instance Sinuany-Stern et al, 1994[6]) where a direct internal comparison of departments of different disciplines belonging to a single university has been performed. The latter approach does not assure comparability between the units of assessment. As demonstrated by Johnes and Taylor (1990)[7], the subject factor is of extreme importance in the assessment of university performance.

5.2.1 The Applicant/ Institution Interface

Different factors affect university selection by prospective students, some 'soft' factors such as atmosphere of the university and reputation are difficult to model, as discussed in Chapter 3, *The Applicant's Perspective.* However, significant input/ output factors such as entry requirements*, teaching and research ratings, employment rates, accommodation availability, library spending, can be used to construct a DEA model for each department of the university (see Figure 5-3), that measures its performance, depending on the preferences of the catchment population for the department. Note that a DEA model is solved for each discipline. The resulting aggregate performance measure (APM), for each department, will reflect the *real* performance of the department in comparison with other departments of the same discipline in other universities, from the perspective of the applicant.

* Attained entry scores rather than entry requirements might be a more suitable measure. However, the latter are widely available, whereas the former are not.

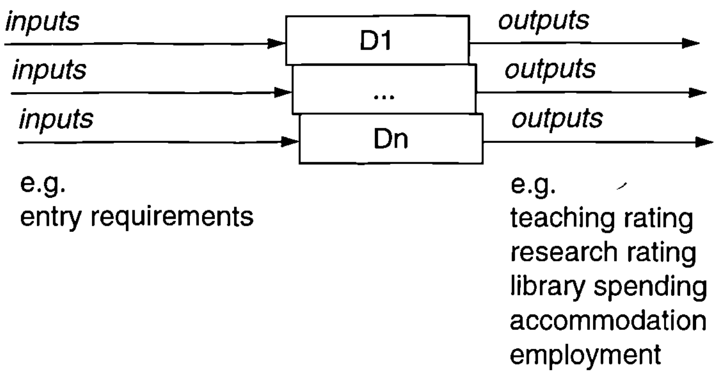

Figure 5-3: A DEA model of *Reality*

On the other hand, the perceived performance, i. e., the *image* of a programme of studies might not coincide with this real performance. The most immediate measure of image can be provided by the entrance requirements for the programme of studies. However, these vary considerably among subjects, and do not give a fair means of comparison between departments. A better measure would be the difference between the entrance requirements for the subject and the national average entrance requirements for that subject. Subjects with a positive *image* value would be above average and those with negative *image* value below average. Perhaps a more appropriate measure for universities that aim to be at the top of the league tables, in terms of quality of applicants, is the distance between the entry requirements of the subject and the national maximum observed entrance requirement for that subject. This measure is also more in accordance with a frontier (best practice) method like DEA. The university will then be able to use this two dimensional analysis of performance, *reality* vs. *image*, to guide the development of the department. In Figure 5-4, four general situations: best practice, improve marketing, query, and review, are possible:

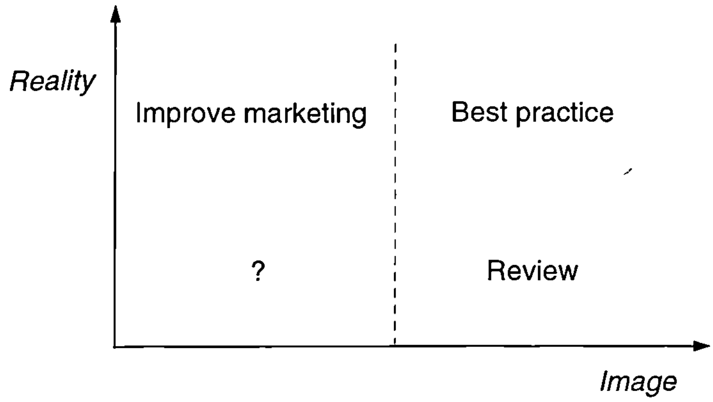

Figure 5-4: *Image* vs. *Reality*

Best practice: The real performance of the department is good and so is the image. This is a benchmark department and its practice should be analysed and emulated by others, where possible.

Improve marketing: The objectively measured reality is good but the image the market has does not accompany it. This department/ programme of studies might well benefit from better marketing and publicity to improve its image and the appropriateness of applicants.

Query: The department is not performing well in both dimensions. This might indicate a genuinely bad performer with little chance of improvement, whose activity should be ceased. Alternatively it may be a new or under-resourced unit that needs investment to achieve its full potential and eventually move towards the best-practice quadrant.

Review: This is a situation that is unlikely to occur, since students and their families are increasingly well informed about higher education. It is an unsustainable position, as bad performance will eventually catch up with an out-of-date good image. In the unlikely case that it does occur, a concerted effort to improve performance is required.

5.2.2 The State/ Institution Interface

The main purpose of the funding councils is to perform the following functions: the distribution of resources for teaching and research; the assessment of the quality of higher education; and the monitoring of the universities financial probity; according to guidelines that are controlled at the centre, both through legislation as well as by the Secretary of State for the sector.

Again, as with the applicant, a DEA model for each discipline (see Figure 5-5) can be constructed, that captures the performance of the department from the perspective of the State. The resulting APM, for each department in relation to the other departments of the same discipline in other universities, will indicate *accountability* to state objectives.

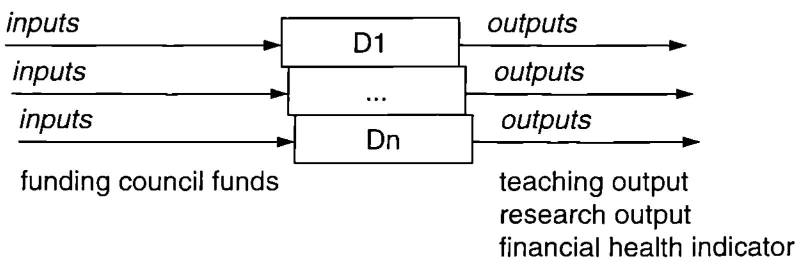

Figure 5-5: A DEA model of *Accountability*

Funding council funds are the money allocated to the department distributed mainly by formula, which takes account of the size and activities of the individual department and the quality of its research[8]. The teaching and research outputs can be considered as the volume of research and teaching moderated by a quality factor, by calculating a measure of volume x quality rating. Teaching volume can be given by the number of FTE students by department. Research volume by the number of FTE academic research active staff (those submitted for the research assessment exercise) by department. The teaching and research quality can be translated by the ratings resulting from the teaching quality assessment and the research assessment exercises. A list of financial indicators that measure the financial health of the institution are published by HESA at institutional level only. The amount of funding council funds are also only available via HESA at institutional level.

In contrast to *accountability*, the pursuit of state objectives, universities will also want to pursue objectives related to their own independence, that might or might not coincide with state ones. This *autonomy* aspect of performance can be measured by the amount of non-funding-council funds the department is able to attract. This internal value added can be measured in terms of earned income, such as from overseas and full-cost-fee students, research councils and charities, other research grants and contracts, services rendered, etc. From the interaction of the two measures of performance, *autonomy* and *accountability*, emerging from the relationship between the State and institution, different possibilities arise as seen in Figure 5-6.

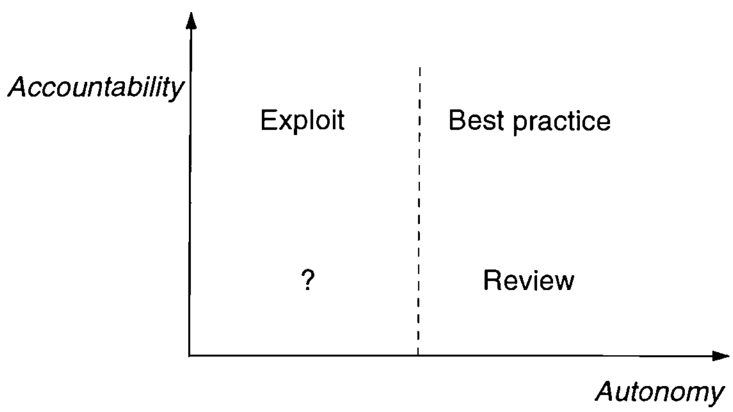

Figure 5-6: *Autonomy* vs. *Accountability*

Best practice: The department is producing good quality teaching and research at the desired level, while keeping financial probity. At the same time it is raising a good amount of income in contribution for the autonomy of the institution. Its practice should be disseminated to other departments, where possible.

Exploit: The department is not exploiting to the full its capacities. Although it is performing well, keeping to its contract with the funding councils, it should be possible to go beyond and engage in other activities that will bring even more resources.

Query: This position represents a question mark that needs to be investigated further. Either it is a unit which no longer is important and closure be considered, or it is a new department that is worth nurturing for its future prospects.

Review: This is either a case that legitimately is pursuing objectives quite different from the State but important for the institution, in which case the situation might be sustainable. Or the department is not keeping to its contract with the State, and needs to improve in the accountability dimension in order to assure basic state funding.

5.2.3 The University 'Sandwiched' Between Different External Evaluations

The juxtaposition of these different dimensions of performance will raise questions for the management of the institution under analysis (see Figure 5-7). How to allocate resources: reward the stars?, help the weak?, is there congruence or conflict between the different perspectives? It is possible that a department will not be in the same quadrant of performance for both internal/ external interfaces, as is the case of department D1 in Figure 5-7.

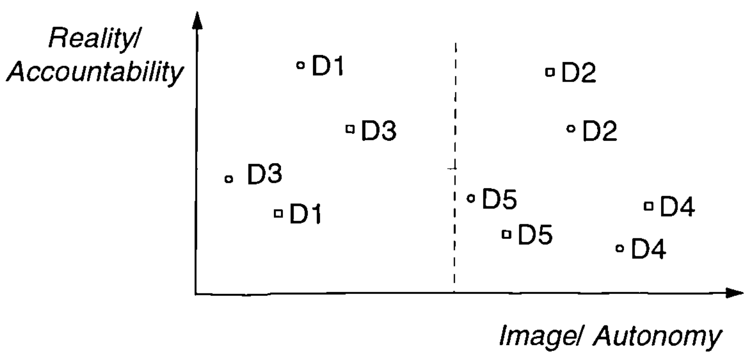

Figure 5-7: PM dimensions: congruence or conflict?

In conclusion, the framework for institutional performance management developed in this section takes into account different levels of stakeholders interested in the performance of universities. It explicitly explores the interaction between external and internal performance measurement at two interfaces: the applicant/ institution, and the State/ institution. Departing from previous studies, that directly compare departments of the same institution; here departments are first compared against their peers in other universities, and only afterwards are their performances compared internally to other departments of the university being studied.

The fact that there is no data publicly available at the required level of disaggregation makes it impossible at present for the analysis of the *accountability* dimension at departmental level. To answer the questions raised by the interaction of the different perspectives on university performance, an implementation of the framework, slightly modified because of the absence of the accountability dimension, is put into practice, and the results discussed in the next section.

5.3 An Application to the University of Warwick

5.3.1 Image vs. Reality

In the following analysis the applicant to the University of Warwick (UW) under consideration will be the traditional 18-year-old undergraduate, directly out of school. The analysis can be extended to other categories of applicants, with different profiles as considered in Chapter 3.

UW collects information from its prospective students during open days. The Warwick applicant is very much aware of the university as being at the top end of league tables, such as the *Times'*[9], where it was placed fifth. Individual determining factors mentioned for choosing Warwick are its success (completion and employment) rates, the quality, reputation and rank compared to other universities, and teaching standards. The atmosphere of its campus, friendliness of staff and students is also mentioned, and UW is aware of this contribution from the high conversion rate from students attending open days. High entry requirements have the 'luxury-item effect' recognised in marketing, the higher the price, the higher the demand. This has happened, for instance, with the mathematics degree, which has seen a dramatic increase in applications, since raising entry requirements to the level of Cambridge, a direct competitor. The course structure is an obvious factor, and distance from home is also mentioned. This is not only related to the so called 'laundry visit', but also with the increasing awareness of student debt, both by the students and their families.

In conclusion, the UW applicant can be characterised as a high academic achiever, who is interested in factors relating to the programme of studies, such as teaching and research quality, and employment prospects; and factors relating to the institution, such as facilities available to students, availability of student accommodation, and living costs. These characteristics can be translated into a DEA model, see Figure 5-8. There will be a similar model for each discipline under consideration. For each model the comparator set of DMUs is based on those departments of the same discipline, which are research active. The reason being that Warwick considers itself as a research-led institution. For a description of the model used, see Appendix 5-1. The analysis was performed for ten subjects of the UW that have already been subjected to assessments of both teaching and research: applied social work, business and management, chemistry, computer science, English, history, law, French, German, and sociology.

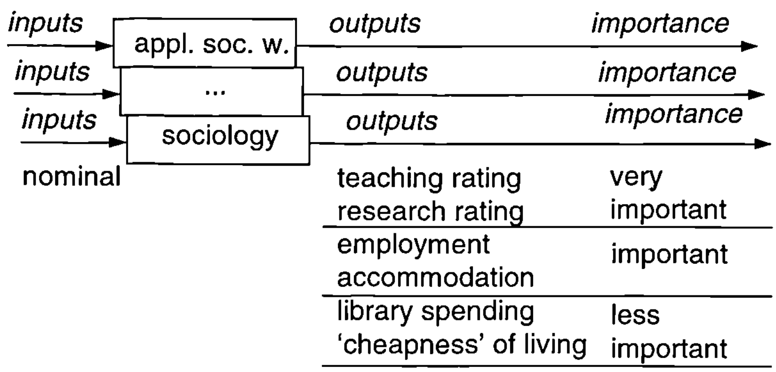

Figure 5-8: A DEA model of *Reality* for the UW

Teaching and research ratings for the departments are the result of the last teaching quality assessment and research assessment exercises undertaken by the funding councils for England, Wales, and Scotland. The teaching score is on a scale of 1 to 24, and the research score on a scale of 1 to 7 (the original scale being 1, 2, 3b, 3a, 4, 5, 5*). The employment score is the proportion of home students from the department taking up employment or going on to further study. The accommodation score represents the proportion of students in residential accommodation maintained by the university. Library spending is an indication of the level of facilities offered to students by the university. Data on teaching and research ratings, accommodation, and library spending is available from the *Times Higher Education Supplement Internet Service*[10]. Data on departmental employment scores was provided by HESA. 'Cheapness' of living is a measure derived from the average student debt, whose data is collected in the PUSH guide[11], and is computed as a 'saving' when considering the most expensive location to study, corresponding to the highest debt. The PUSH guide uses debt as a proxy for cost of living. However, it should be borne in mind that debt might be more dependent on the kind of personal circumstances incurred by the students attending a particular institution, rather than the cost of living in the area.

All indicators comply to the rule 'the more the better', which will allow the use of an output oriented model, with the sum of the virtual weights of all the factors being one.

This property will enable the intuitive application of virtual weights restrictions that translate the different importance of different factors (see Appendix 5-1). Since all factors included in the model are deemed important for the Warwick applicant, they are made to have a minimum contribution of 5%. Important factors are then weighted at least the double of less important factors, and very important factors at least the double of important ones. In general, it is for the decision maker to decide on the importance of the different factors. In the current application the valuation of the factors was inferred from the analysis of data concerning the typical UW applicant, thus for a different university the valuation would be different. To be completely accurate a survey of the order of importance for the factors should have been made for all departments, as it is natural to expect that the characteristics of the applicants will vary from subject to subject. Unfortunately data was not available at departmental level. The results for the dimension of *reality* for the subjects under analysis are given in Table 5-2 to Table 5-11, Appendix 5-2. Image was calculated as the difference between the average entry requirements of the subject and the national maximum observed average entrance requirement for that subject. The A-level entry points are calculated attributing 10 points to grade A, 8 to grade B, etc. The maximum possible score is 30, however, in some subjects, the observed maximum is below this.

The resulting matrix of performance is as in Figure 5-9. The values that are used for the axes that delineate the quadrants should be decided judgementally according to what is thought to be good performance by the organisation under analysis.

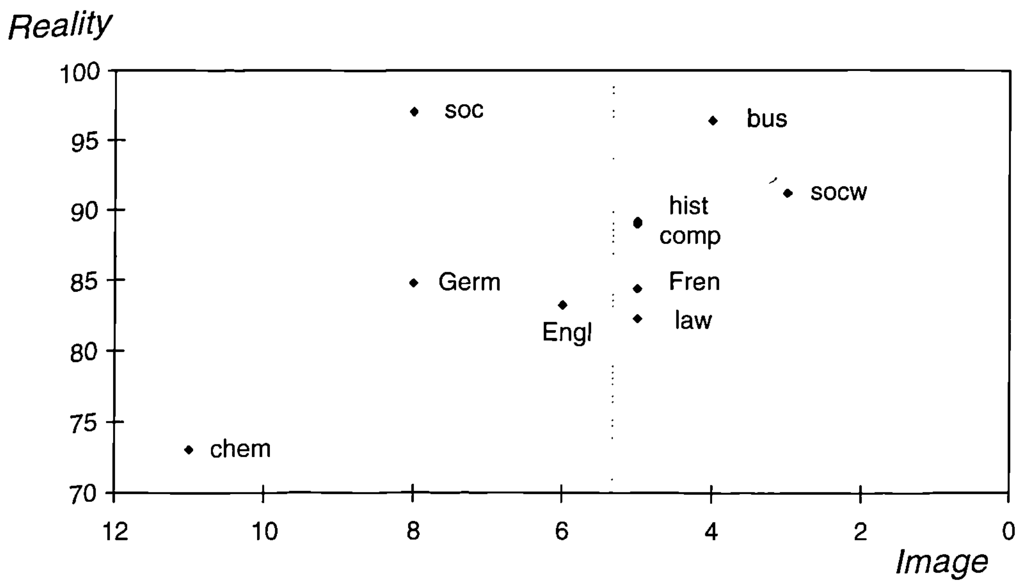

Figure 5-9: *Image* vs. *Reality* matrix at the UW

All subjects at UW are in the top 5 of the *reality* league with APM scores above 80%, when compared to departments of the same discipline in other universities, except for chemistry (see Appendix 5-2). Despite this there are significant differences on how a department can still improve in relation to the best performers of its discipline. For instance, business and French achieve the same ranking, when compared with their peers, but there is significantly more scope for improvement in French with an APM of 84% than business studies with an APM of 96%.

As to *image*, again most departments perform well in relation to their peers, with the expected exception of chemistry. German, English and especially sociology seem to have an *image* that is poorer than what they have in *reality* to offer. These departments may well benefit from marketing their strengths. Chemistry, the most worrying case, was the subject of the dilemma close/ invest after the outcome of the 1996 RAE, the result of which was to invest (more senior positions were allocated to the department), since it was considered a discipline of strategic importance to the university. (A strategic objective of UW is developing the science and technology base of the University).

5.3.2 Autonomy vs. Accountability

Given that funding council funds are not publicly available at the desired level of disaggregation, the measurement of *accountability*, at present, can only be calculated when the unit of assessment is the institution, and not the department. This analysis and the comparison of institutions from the State perspective was dealt with in Chapter 4.

At the moment, teaching ratings will be available by 63 subjects of study. Up to September 1996, the assessment of 23 subjects had been completed. The completion of the first full cycle of subject assessments will only be finished by December 2001. Research ratings are the result of the Research Assessment Exercise (RAE), which has been performed 4 times, since 1986. In the RAE96 the ratings were produced for 69 different units of assessment (UOA), which might or might not coincide with the subject categories used for the teaching quality assessment exercise. Data on students, staff, and expenditure by 35 academic cost centres (ACC) has been compiled, for the first time, for the academic year 1995/96[12] for English institutions by HEFCE. Since then, HEFCE proposed extending the number of academic cost centres to 39[13], and later on the number of ACC were increased to 40[14].

When teaching assessment of all disciplines has been completed, and data on the academic cost centres made available, then the methodology proposed could, in principle, be used to compare all academic cost centres, assuming that a mapping of the teaching and research ratings to these ACC is possible.

The dimension of *autonomy*, however, can be studied at departmental level by the institution. In Figure 5-10 four measures that contribute to autonomy are displayed for the ten departments under analysis. They represent the contribution of the department to university income from overseas-student fees, research grants, research contracts, and other earned income, in pounds per full-time-equivalent staff for the academic year 1994/95. Only overseas-student fees and other earned income are truly discretionary income that can be used on other projects that the university is interested in. Most income from research grants and contracts is spent on the projects they were meant to fund anyway. However, they are included because they fund activities that the departments are interested in undertaking, beyond the normal activities funded by the funding councils. They represent the pursuit of departments' own interests and project their activities to the larger academic and industrial community, bringing recognition and possibly more resources. Even the funding councils recognise this effect, taking account of this external research income in their funding formula, rewarding the ability of the department in attracting funding from alternative sources.

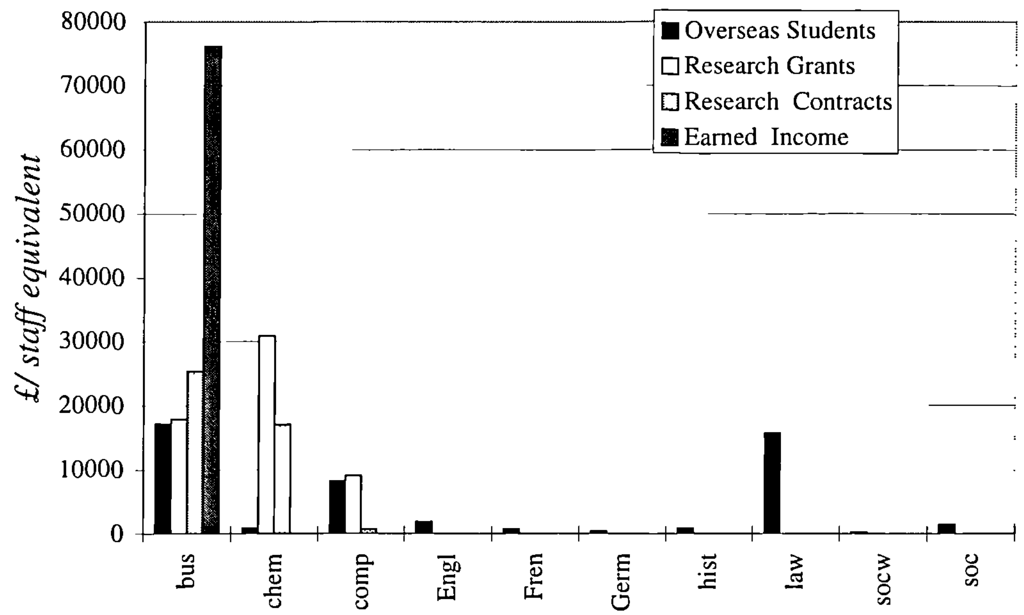

Figure 5-10: 'Absolute' *Autonomy* at the UW

As with the other dimensions of performance, ideally the measurement of *autonomy* should be in relation to the department's peers, as it is not legitimate to compare the amount of earned income obtained by a business studies department and an English department, for example. However, once again, the information is not available at the level of disaggregation required for all the factors considered. Information is only available for external (non-funding councils) research income attracted by each department, obtainable from the RAE96 database[15]. Figure 5-11 shows the external income attracted by UW departments during the three academic years (1992/93 - 1994/95) preceding the RAE96 per FTE academic staff, in relation to the average and observed maximum for that subject, which is more in line with the framework developed.

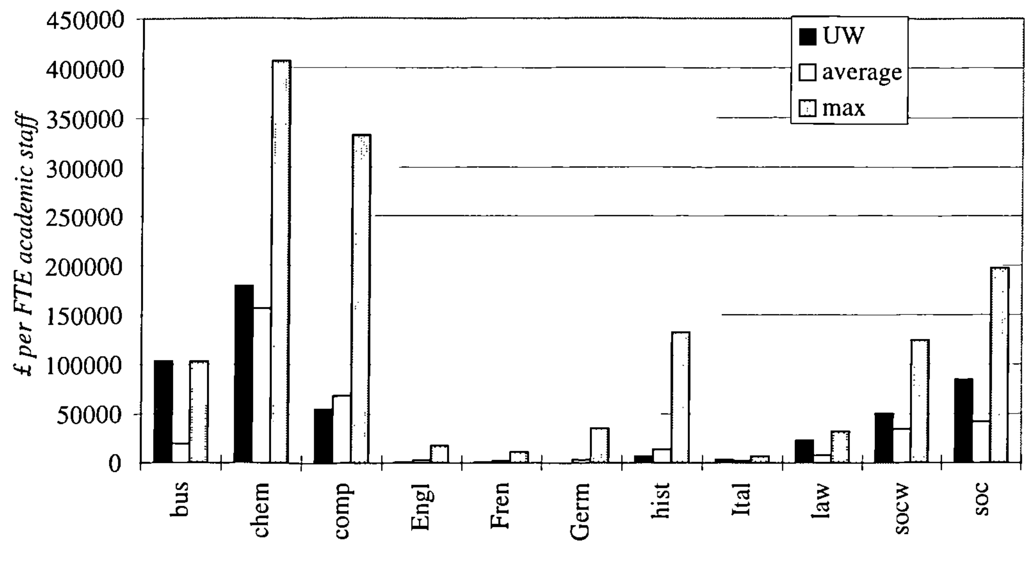

Figure 5-11: 'Relative' *Autonomy* at the UW

Business studies at UW is the benchmark for the subject in all universities, as it achieves the observed maximum. Chemistry, law, social work, and sociology are all above average, although, except for law, the situation can still be significantly improved. History has a lot of scope for improvement in this measure. As for the languages, English, French, German, and Italian, the situation is identical for all the sector, which shows difficulty in attracting external research income.

5.3.3 The Whole Picture

Although, the juxtaposition of the two interfaces as in Figure 5-7 cannot be implemented, because of the lack of information on the *accountability* dimension, a visual compact picture of the interaction of the three remaining dimensions: *image*, *reality*, and *autonomy*, is presented in Figure 5-12. The Reality dimension is measured by the DEA score of the Reality model (see Figure 5-8). Image is measured by the difference between the average entry requirement of the subject and the national maximum observed for that subject. Autonomy, as for Image, is the difference between the income attracted by each department and the national maximum observed for that subject.

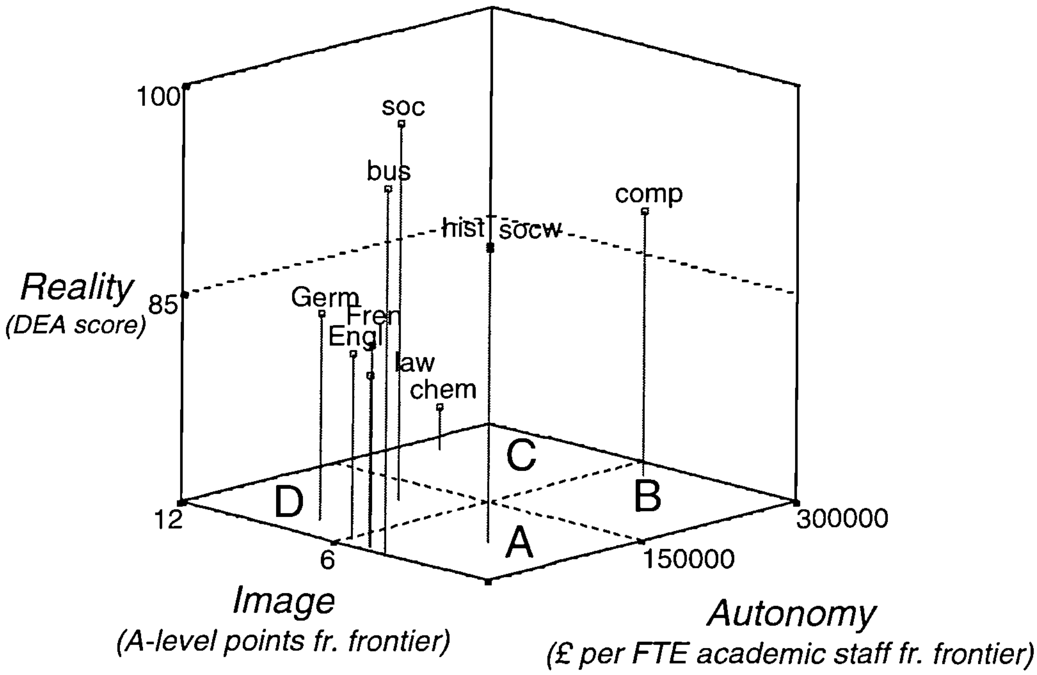

Figure 5-12: Image vs. Reality vs. Autonomy

The positioning of the disciplines on the floor of the picture, represents their performance on the image and autonomy dimensions, over which UW can probably change the situation more quickly than in the reality dimension. Quadrant A being the best practice quadrant (nearer the frontiers of the two dimensions), and C representing weak performance on both dimensions. B represents good performance on the image dimension, and not so go in the autonomy dimension. In D are located the disciplines with good performance on the autonomy dimension, and not so good on the image dimension. The height of the spikes for each discipline represent how well they perform on the reality dimension. Clearly business studies, sociology, and computer science are best performers in the reality dimension, above the 85% APM line. Six categories emerge from the analysis of Figure 5-12, summarised in Table 5-1.

Table 5-1: Summary of global performance

	Reality	
Image x Autonomy	*< 85%*	*> 85%*
A (+, +)	Engl Fren law hist socw	bus
B (+, -)	-	comp
C (-, -)	chem	-
D (-, +)	Germ	soc

Business studies is the star discipline with very good performance in all dimensions. English, French, law, history, and social work perform very well in the image and autonomy dimensions, but could still significantly improve the reality dimension. Computer science is quite well positioned on the image and reality dimension, but could further exploit the autonomy dimension. German and sociology can substantially increase their image value in the market for undergraduates, but both perform well in the autonomy dimension. German can additionally increase its performance in the reality dimension, whereas sociology is already quite well positioned relatively to its peers in other universities. Chemistry is clearly lagging, having disappointing performances in all dimensions.

In conclusion, the application of the framework for institutional performance management to the UW, although not complete for lack of data, has shown to give some insights into the performance of different disciplines within UW. The interaction of three different dimensions of performance, reveal different groupings, for which different strategies are required. Best performers, such as business studies can be used as benchmark departments, whose practice can be emulated by others performing less well. Perhaps the easiest situation to address is when the reality is good, and only better marketing for the department or more exploitation of the autonomy dimension is required. On the other hand, bad performance on the reality dimension will probably take more time and resources to address, maybe through recruitment of new staff and/ or targeted investment from central management. Ultimately, decisions on allocation of resources, flux of funds to reward good performers, increase funds to units to improve their performance, or even closure of units, will remain with the decision makers.

5.4 Concluding Remarks

This chapter presents a novel approach for the comparison of departments of different subjects within the same institution. Departing from previous studies, which compare different subject departments directly, in this approach departments are firstly judged against those of the same discipline in <u>other</u> universities to ensure comparability; and then their performance, thus assessed, compared to the other departments of the <u>same</u> university.

Different classes of stakeholders have different motivations to measure performance. Institutions will be affected on the one hand by the State evaluation of them, and on the other by the applicant's. Institutions can benefit from the information provided by matrices of performance that take into account these two interfaces between the institution and the applicant on the one hand, and the institution and the State on the other. The methodology proposed enables the institution to allocate resources among departments in a more informed way and thus enhance its performance as a whole.

The application of the methodology to the University of Warwick faced the constraints of data availability. Information on performance is vital to enhance performance, but regrettably universities have not been able to agree on a common way of measuring this performance. This stems from the fact that universities have been systematically submitted to evaluations that do not take into account their mission and objectives. The problem has been aggravated since the 1992 merging of the polytechnic with the university sector, bringing more heterogeneity to the university sector. It seems that all parties involved would profit from having information readily available at the required level of disaggregation (most performance measurement is only meaningful at subject level), so that they could use the information to guide their choices in a admittedly diverse new system of higher education.

Appendix 5-1: DEA model

As the factors under consideration were principally used as outputs, model M4 (output oriented CRS) was used in the analysis, so that the aggregate of the virtual outputs (the product of the output level and optimal weight for that output) summed to 1 (100%) for each unit.

For the reality dimension the following restrictions are added to M4 to express the preference structure of the traditional undergraduate UW applicant:

$u_V y_{Vj_0} \geq 2u_I y_{Ij_0}$ V denotes a very important factor

$u_I y_{Ij_0} \geq 2u_L y_{Lj_0}$ I denotes an important factor

$u_L y_{Lj_0} \geq 0.05$ L denotes a less important factor

Appendix 5-2: Results for *Reality* for UW departments

Table 5-2: Results for business - top of table

Unit	*APM*	*Rank*
LANCASTER	100	1
OXFORD	98.01	2
CAMBRIDGE	96.49	3
WARWICK	**96.42**	**4**
UMIST	95.03	5
ST_ANDREWS	88.67	6
LSE	88.05	7
READING	87.41	8
IMPERIAL	87.05	9
SOUTHAMPTON	86.78	10

Table 5-3: Results for chemistry - top of table

Unit	*APM*	*Rank*
OXFORD	100	1
CAMBRIDGE	99.7	2
ST_ANDREWS	87.32	3
DURHAM	85.66	4
BRISTOL	82.17	5
HULL	81.73	6
SOUTHAMPTON	81.48	7
IMPERIAL	81.03	8
NOTTINGHAM	80.5	9
MANCHESTER	80.1	10
YORK	79.84	11
LEICESTER	78.63	12
BIRMINGHAM	76.34	13
EDINBURGH	76.29	14
SUSSEX	75.38	15
READING	75.11	16
SHEFFIELD	74.88	17
CARDIFF	74.78	18
UMIST	73.89	19
EXETER	73.88	20
NEWCASTLE	73.28	21
WARWICK	**73.08**	**22**

Table 5-4: Results for computer science - top of table

Unit	*APM*	*Rank*
CAMBRIDGE	100	1
OXFORD	99.13	2
YORK	93.19	3
WARWICK	**89.04**	**4**
ST_ANDREWS	83.73	5
MANCHESTER	82.81	6
IMPERIAL	81.64	7
SOUTHAMPTON	81.24	8
KENT	80.68	9
EXETER	79.84	10

Table 5-5: Results for English - top of the table

Unit	*APM*	*Rank*
OXFORD	100	1
CAMBRIDGE	99.03	2
YORK	88.7	3
ST_ANDREWS	86.13	4
WARWICK	**83.28**	**5**
DURHAM	82.44	6
LANCASTER	81.39	7
SUSSEX	80.75	8
BIRMINGHAM	80.24	9
BRISTOL	79.88	10

Table 5-6: Results for French - top of the table

Unit	*APM*	*Rank*
CAMBRIDGE	100	1
OXFORD	97.57	2
ST_ANDREWS	86.61	3
WARWICK	**84.44**	**4**
DURHAM	82.22	5
READING	82.17	6
NEWCASTLE	80.49	7
SUSSEX	80.14	8
HULL	79.81	9
ROYAL_HOLLOWAY	79.77	10

Table 5-7: Results for German - top of the table

Unit	*APM*	*Rank*
CAMBRIDGE	100	1
OXFORD	97.68	2
ST_ANDREWS	86.8	3
WARWICK	**84.82**	**4**
EXETER	83.35	5
NOTTINGHAM	82.24	6
SWANSEA	82.16	7
MANCHESTER	81.33	8
DURHAM	78.58	9
LEICESTER	77.27	10

Table 5-8: Results for history - top of the table

Unit	*APM*	*Rank*
OXFORD	100	1
CAMBRIDGE	98.11	2
ST_ANDREWS	90.73	3
WARWICK	**89.23**	**4**
YORK	85.45	5
DURHAM	84.96	6
LSE	84.62	7
HULL	84.16	8
LANCASTER	81.79	9
ROYAL_HOLLOWAY	80.76	10

Table 5-9: Results for law - top of the table

Unit	*APM*	*Rank*
OXFORD	100	1
CAMBRIDGE	98.65	2
MANCHESTER	82.74	3
DURHAM	82.36	4
WARWICK	**82.33**	**5**
LSE	81.59	6
SHEFFIELD	80.66	7
NOTTINGHAM	80.22	8
BRISTOL	80.02	9
ESSEX	79.35	10

Table 5-10: Results for social work - top of the table

Unit	*APM*	*Rank*
YORK	100	1
STIRLING	100	1
LANCASTER	95.27	3
WARWICK	**91.28**	**4**
KEELE	91.24	5
EAST_ANGLIA	89.17	6
BRISTOL	86.06	7
HULL	85.44	8
SOUTHAMPTON	83.01	9
SWANSEA	82.64	10

Table 5-11: Results for sociology - top of the table

Unit	*APM*	*Rank*
CAMBRIDGE	100	1
LANCASTER	98.53	2
WARWICK	**97.05**	**3**
YORK	94.59	4
ESSEX	94.3	5
OXFORD	94	6
LOUGHBOROUGH	91.37	7
MANCHESTER	90.87	8
SURREY	89.62	9
SUSSEX	85.68	10

5.5 References

[1] Dyson RG, Athanassopoulos AD, Thanassoulis E (1994). Performance measurement systems, managerial control and data envelopment analysis. Research Paper 117, Warwick Business School.

[2] Hax AC and Majluf, NS (1990). The use of growth-share matrix in strategic planning. In: Dyson RG (ed). *Strategic Planning: Models and Analytical Techniques*. Wiley: Chichester.

[3] Dyson RG, Thanassoulis E and Boussofiane A (1990). Data envelopment analysis. In: Hendry LC and Eglese RW (eds). *Tutorial Papers in Operational Research.* Operational Research Society: Birmingham.

[4] Boussofiane A, Dyson RG, Thanassoulis E (1991). Applied data envelopment analysis. *Eur J Opl Res* **52**: 1-15.

[5] Athanassopoulos AD and Thanassoulis E (1995). Separating market efficiency from profitability and its implications for planning. *J Opl Res Soc* **46**, 20-34.

[6] Sinuany-Stern Z, Mehrez A, and Barboy A (1994). Academic departments efficiency via DEA. *Comput & Ops Res* **21:** 543-556.

[7] Johnes J and Taylor J (1990), *Performance Indicators in Higher Education.* SRHE & Open University Press: Buckingham.

[8] Higher Education Funding Council for England (1996). *A Guide to Funding Higher Education 1996-97, M 10/96.* HEFCE: Bristol.

[9] O'Leary J (1997). *The Times Good University Guide 1997.* Times Books: London.

[10] Times Higher Education Supplement Internet Service (1997). *Research and Teaching Quality Listings.* Times Newspapers Ltd: London.

[11] Rich J and Footman T (eds) (1995). *The PUSH Guide to Which University 96.* PUSH in association with Lloyds Bank: London.

[12] Higher Education Funding Council for England (1997). *Management Statistics Produced from 1995-96 HESA Records, CL 13/97.* HEFCE: Bristol.

[13] Higher Education Funding Council for England (1997). *HEFCE Review of Academic Cost Centres for 1998-99, CL 20/97.* HEFCE: Bristol.

[14] Higher Education Funding Council for England (1997). *Assigning Departments to Academic Cost Request, 97/25*. HEFCE: Bristol.

[15] Higher Education Funding Council for England (HEFCE), the Scottish Higher Education Funding Council (SHEFC), the Higher Education Funding Council for Wales (HEFCW) and the Department of Education for Northern Ireland (DENI) (1996). *1996 Research Assessment Exercise: The Outcome, RAE96 1/96*. HEFCE: Bristol.

6. THE USE OF VIRTUAL WEIGHTS RESTRICTIONS IN DATA ENVELOPMENT ANALYSIS

6.1 Introduction

The application of DEA to concrete situations has motivated the use of weights restrictions to curb the complete freedom of variation of weights allowed by the original DEA model. The problem of allowing total flexibility of the weights, is that the values of the weights obtained by solving the unrestricted DEA program are often in contradiction to prior views or additional available information.

Thompson et al. (1986)[1] were the first authors to propose the use of weights restrictions to increase discrimination of the results of a DEA problem to support the siting of a laboratory, where only six alternatives were under consideration. Their technique included the imposition of acceptable bounds on ratios of multipliers, to solve a choice problem. Dyson and Thanassoulis (1988)[2] were concerned about the omission of particular inputs or outputs from the efficiency score, through the assignment of zero weights. They suggested imposing meaningful bounds directly on individual multipliers based on average input levels per unit of output. Charnes et al. (1990)[3], in another approach to the problem, suggested transforming input-output data to simulate weights restrictions, where DMUs are assessed on the basis of the input-output levels of pre-selected DMUs which were a priori recognised by experts as being efficient.

One of the problems with directly restricting the multipliers, i.e. *absolute* weights restrictions, is that they are dependent on the units of measurement of the inputs and outputs. Virtual input/ output, the product of the input/ output level and optimal weight for that input/ output, however, is dimensionless. The virtual inputs and outputs of a DMU reveal the relative contribution of each input and output to its efficiency rating. The higher the level of virtual input/ output, the more important that input/ output is in the efficiency rating of the DMU concerned. Therefore use of virtual inputs and outputs can help to identify strong and weak areas of performance. In order to avoid the problem of *absolute* weights restrictions, Wong and Beasley (1990)[4] proposed the use of *virtual* weights restrictions, and in particular, the use of *proportional* virtual weights restrictions, which were intended to make it easier for the decision maker to quantify value judgements in terms of percentage values. Because of the intuitive approach to setting up value judgements, virtual weights restrictions have been widely used in this thesis.

A comprehensive review of the evolution, development and research directions on the use of weights restrictions can be found in Allen et al. (1997)[5]. In this review the consequences for the interpretation of the results from DEA models with weights restrictions has been analysed for *absolute* weights restrictions. The analysis of the pros and cons of the use of *virtual* weights restrictions is proposed as a further direction of research. This chapter proposes to contribute to that analysis.

The intention of incorporating value judgements might be, as seen above, to incorporate prior views or information regarding the assessment of efficient DMUs. On the other hand, there might be two levels of decision-making, the DMU (for instance, a department or university), and the corporate top management or external evaluator (for instance the State, or the applicant in relation to a university or department). The DMU might use its value judgements if it wants to use the assessment for benchmarking itself against other DMUs (as in Chapter 5, *The Institutional Perspective*). However, if the evaluation is done by an external agent, the expressions *DMU* and *decision maker* might be misleading, as the decision maker is, in fact, at a different level. In this case, the DEA assessment becomes a game between the *unit of assessment* (UOA) trying to show itself in the best possible light, and a higher level decision maker, i.e. a *controller* imposing its preference structure. In Chapter 3 and Chapter 4 the institution or department is the UOA with the applicant or the State being the controller. A similar situation occurs in the regulated industries where the regulator is the controller.

Allen et al. (1997)[5] point out that the substantial changes to the UOA's current mix of input and output levels indicated by the imposition of weights restrictions might be beneficial. It might lead to the conclusion that the current mix is inadequate given the controller's preferences. The same goes for the deterioration of current levels of some inputs and outputs.

In the next sections different methods of setting up virtual weights restrictions are discussed, and their pros and cons, in terms of their effect on the efficiency value, target setting, and peer comparators.

6.2 The Use of *Proportional* Virtual Weights Restrictions

The method developed by Wong and Beasley (1990)[4] to restrict flexibility in DEA is based upon the use of proportions. Conceptually the proportional virtual output r of UOA j represents the importance attached to the output measure (a similar reasoning can be applied to an input factor). Thus the controller can set limits on this proportion to reflect value judgements, as follows:

$$a_r \leq \frac{y_{rj}u_r}{\sum_{r=1}^{s} y_{rj}u_r} \leq b_r \text{ for an output factor, and}$$

$$c_i \leq \frac{x_{ij}v_i}{\sum_{i=1}^{m} x_{ij}v_i} \leq d_i \text{ for an input factor.}$$

Note that these kinds of restrictions are UOA specific, which raises questions for their implementation, namely to which UOAs should the restrictions apply, and what effect they have on other UOAs. The authors suggest the following three different alternatives. Since the models in this thesis have an output orientation, an output oriented model (i.e. where $\sum_{r=1}^{s} y_{rj_0}u_r = 1$) is considered to explore the consequences of the use of the different approaches. The use of an input oriented model would lead to similar conclusions.

6.2.1 Proportional virtual weights restrictions apply only *to the target UOA* j_0.

When proportional weights restrictions are applied to the target UOA j_0, each UOA is assessed with two additional constraints for each factor (output or input) being restricted:

$$a_r \le \frac{y_{rj_0} u_r}{\sum_{r=1}^{s} y_{rj_0} u_r} \le b_r, \forall r \text{ and } c_i \le \frac{x_{ij_0} v_i}{\sum_{i=1}^{m} x_{ij_0} v_i} \le d_i, \forall i$$

Note that, so far as the target unit is concerned, the above constraints reduce to:

$$a_r \le y_{rj_0} u_r \le b_r, \forall r \text{ and } c_i \le \frac{x_{ij_0} v_i}{\sum_{i=1}^{m} x_{ij_0} v_i} \le d_i, \forall i$$

Adding the weights restrictions to the multipliers formulation, it becomes:

$$\begin{array}{llccl}
\frac{1}{e_0} = \min \sum_{i=1}^{m} x_{ij_0} v_i & & & & \\
s.t. & & & & \\
 & \sum_{r=1}^{s} y_{rj_0} u_r & = & 1 & \\
\sum_{i=1}^{m} x_{ij} v_i & -\sum_{r=1}^{s} y_{rj} u_r & \ge & 0, & \forall j \\
 & y_{rj_0} u_r & \ge & a_r, & \forall r \\
 & -y_{rj_0} u_r & \ge & -b_r, & \forall r \\
x_{ij_0} v_i - c_i \sum_{i=1}^{m} x_{ij_0} v_i & & \ge & 0, & \forall i \\
-x_{ij_0} v_i + d_i \sum_{i=1}^{m} x_{ij_0} v_i & & \ge & 0, & \forall i \\
v_i & & \ge & \varepsilon, & \forall i \\
 & u_r & \ge & \varepsilon, & \forall r
\end{array}$$

Since these kinds of restrictions are UOA specific, as they are dependent on the target unit's input/ output levels, the target unit will thus be 'imposing' different, and possibly unreasonable, restrictions on the virtuals of the other units when assessing the target unit. The virtual output r for UOA j_0 will be $y_{rj_0}u_r$. For a unit j_l with output level inferior to that of unit j_0, the virtual output will be constrained to be $a_r \le y_{rj_l}u_r \le y_{rj_0}u_r \le b_r$. For a unit j_m with output level superior to that of unit j_0, the virtual output will be constrained to be $a_r \le y_{rj_0}u_r \le y_{rj_m}u_r \le b_r$. In all cases the allowable virtuals for a unit $j \ne j_0$ will be more restrictive than for the target unit, i.e. the assessment is more 'lenient' to the target UOA.

The consequences for the measure of efficiency and target setting can be better appreciated from the envelopment formulation:

$$\frac{1}{e_0} = \max \theta_0 + \sum_{r=1}^{s} a_r\alpha_r - \sum_{r=1}^{s} b_r\beta_r + \varepsilon\sum_{i=1}^{m} s_i^- + \varepsilon\sum_{r=1}^{s} s_r^+$$

$$\begin{array}{lllllll}
s.t. & & & & & & \\
 & \sum_{j=1}^{n} x_{ij}\lambda_j & +x_{ij_0}\gamma_i - x_{ij_0}\sum_{i=1}^{m} c_i\gamma_i & & & & \\
 & & -x_{ij_0}\delta_i + x_{ij_0}\sum_{i=1}^{m} d_i\delta_i \quad +s_i^- & & = & x_{ij_0}, & \forall i \\
y_{rj_0}\theta_0 & -\sum_{j=1}^{n} y_{rj}\lambda_j & +y_{rj_0}\alpha_r - y_{rj_0}\beta_r & +s_r^+ & = & 0, & \forall r \\
\theta_0 & & & & & \text{free} & \\
 & \lambda_j & & & \ge & 0, & \forall j \\
 & & \gamma_i, \delta_i, \quad s_i^- & & \ge & 0, & \forall i \\
 & & \alpha_r, \beta_r, & s_r^+ & \ge & 0, & \forall r
\end{array}$$

The targets for UOA j_0 will thus be:

$$\sum_{j=1}^{n} x_{ij}\lambda_j^* = x_{ij_0}\left(1 - \gamma_i^* + \sum_{i=1}^{m} c_i\gamma_i^* + \delta_i^* - \sum_{i=1}^{m} d_i\delta_i^*\right) - s_i^{-*}, \forall i$$

$$\sum_{j=1}^{n} y_{rj}\lambda_j^* = y_{rj_0}\left(\theta_0^* + \alpha_r^* - \beta_r^*\right) + s_r^{+*}, \forall r$$

Two new variables α_r and β_r appear in the objective function of the linear program, which might not be negligible in comparison with θ_0. The efficiency measure, e_0^* will no longer necessarily approximate the inverted radial expansion factor $1/\theta_0^*$, as it would in an unrestricted model. Moreover this measure will depend on the limits imposed for the proportional virtual outputs in an output oriented model (and on the limits imposed for the proportional virtual inputs in an input oriented model).

As to the targets, two different situations arise for inputs and outputs. For both input and output factors, the targets obtained can mean either an improvement or a deterioration in the current level. However, for an input factor, its target will also depend on the limits imposed (c_i and d_i) for all the proportional virtual inputs. Because of this characteristic, it would seem more appropriate to impose this kind of weights restriction only on the virtual outputs for an output oriented model, and only on the virtual inputs for a virtual input model. This procedure would be more in line with 'classical' DEA, where in the input oriented models, one focuses on maximal movement towards the frontier through proportional reduction of inputs, whereas in the output oriented models, one focuses on maximal movement via proportional augmentation of outputs, but not both simultaneously. The choice of model will depend on which factors under consideration are more easily controlled by the UOA.

In conclusion, this alternative makes the assessment of performance more 'lenient' to the target UOA, in comparison with all other UOAs. Also, in an output oriented model, the restrictions applied to the target unit for the outputs affect the efficiency score and output targets, whereas the restrictions applied to the inputs will affect the input targets in an unconventional way.

6.2.2 Proportional virtual weights restrictions apply to all *UOAs* j

As seen in the previous section, applying proportional virtual weights restrictions only to the target UOA actually biases the analysis in favour of the target unit, as the other units have to endure tougher restrictions than the target UOA. The second approach, however, proposes that all virtual weights restrictions should apply to all UOAs:

$$a_r \leq \frac{y_{rj}u_r}{\sum_{r=1}^{s} y_{rj}u_r} \leq b_r, \forall r, j \text{ and } c_i \leq \frac{x_{ij}v_i}{\sum_{i=1}^{m} x_{ij}v_i} \leq d_i, \forall i, j$$

This approach is computationally expensive, as for each input or output factor being constrained, $2n$ inequalities are added to each linear program.

Adding the weights restrictions to the multipliers formulation, it becomes:

$$
\begin{array}{llll}
\dfrac{1}{e_0} = \min \sum_{i=1}^{m} x_{ij_0} v_i & & & \\
s.t. & & & \\
\sum_{r=1}^{s} y_{rj_0} u_r & = & 1 & \\
\sum_{i=1}^{m} x_{ij} v_i - \sum_{r=1}^{s} y_{rj} u_r & \geq & 0, & \forall j \\
y_{rj_0} u_r & \geq & a_r, & \forall r \\
-y_{rj_0} u_r & \geq & -b_r, & \forall r \\
y_{rj} u_r - a_r \sum_{r=1}^{s} y_{rj} u_r & \geq & 0, & \forall r, j \neq j_0 \\
-y_{rj} u_r + b_r \sum_{r=1}^{s} y_{rj} u_r & \geq & 0, & \forall r, j \neq j_0 \\
x_{ij} v_i - c_i \sum_{i=1}^{m} x_{ij} v_i & \geq & 0, & \forall i, j \\
-x_{ij} v_i + d_i \sum_{i=1}^{m} x_{ij} v_i & \geq & 0, & \forall i, j \\
v_i & \geq & \varepsilon, & \forall i \\
u_r & \geq & \varepsilon, & \forall r
\end{array}
$$

The envelopment formulation becomes:

$$\frac{1}{e_0} = \max \theta_0 + \sum_{r=1}^{s} a_r \alpha_{r_0} - \sum_{r=1}^{s} b_r \beta_{r_0} + \varepsilon \sum_{i=1}^{m} s_i^- + \varepsilon \sum_{r=1}^{s} s_r^+$$

$$s.t.$$

$$\sum_{j=1}^{n} x_{ij}\lambda_j + \sum_{j=1}^{n}\left[x_{ij}\left(\gamma_{ij} - \sum_{i=1}^{m} c_i \gamma_{ij} - \delta_{ij} + \sum_{i=1}^{m} d_i \delta_{ij} \right)\right] + s_i^- = x_{ij_0}, \quad \forall i$$

$$y_{rj_0}\theta_0 - \sum_{j=1}^{n} y_{rj}\lambda_j + y_{rj_0}\alpha_{r_0} - y_{rj_0}\beta_{r_0} + \sum_{\substack{j=1 \\ j \neq j_0}}^{n}\left[y_{rj}\left(\alpha_{rj} - \sum_{r=1}^{s} a_r \alpha_{rj} - \beta_{rj} + \sum_{r=1}^{s} b_r \beta_{rj} \right)\right] + s_r^+ = 0, \quad \forall r$$

$$\theta_0 \quad \text{free}$$

$$\lambda_j \geq 0, \quad \forall j$$

$$\gamma_{ij}, \delta_{ij} \geq 0, \quad \forall i, j$$

$$s_i^- \geq 0, \quad \forall i$$

$$\alpha_{rj}, \beta_{rj} \geq 0, \quad \forall r, j$$

$$s_r^+ \geq 0, \quad \forall r$$

The targets:

$$\sum_{j=1}^{n} x_{ij}\lambda_j^* = x_{ij_0} - \sum_{j=1}^{n}\left[x_{ij}\left(\gamma_{ij}^* - \sum_{i=1}^{m} c_i \gamma_{ij}^* - \delta_{ij}^* + \sum_{i=1}^{m} d_i \delta_{ij}^* \right)\right] - s_i^{-*}, \forall i$$

$$\sum_{j=1}^{n} y_{rj}\lambda_j^* = y_{rj_0}\left(\theta_0^* + \alpha_{r_0}^* - \beta_{r_0}^*\right) + \sum_{\substack{j=1 \\ j \neq j_0}}^{n}\left[y_{rj}\left(\alpha_{rj}^* - \sum_{r=1}^{s} a_r \alpha_{rj}^* - \beta_{rj}^* + \sum_{r=1}^{s} b_r \beta_{rj}^* \right)\right] + s_r^{+*}, \forall r$$

As with the previous approach the efficiency score e_0^* will no longer necessarily approximate the inverted radial expansion factor $1/\theta_0^*$. Moreover, the interpretation of the targets becomes increasingly difficult. The targets for a factor (either input or output) become dependent, not only on the value of that factor for the target unit, but also on the value of that factor for all the other units. Additionally, they become dependent on the limits imposed for the virtuals (a_r and b_r for outputs and c_i and d_i for inputs) of all other factors.

The argument that this alternative is computationally expensive is probably the least of its caveats as software and hardware become increasingly more powerful. This approach applies each unit's restrictions to all the other units, and is thus applying the 'worst case' to each unit. As a consequence, and more worrying is the propensity of this approach to lead to infeasible linear programs, as will be shown later in this chapter.

6.2.3 Proportional virtual weights restrictions apply to target and *to an 'average' artificial UOA* j_a

In order to keep to the spirit of their second approach (6.2.2), but to avoid its computational problems, the authors favour a third approach involving an artificial 'average' UOA. The logic behind it is that the proportional weights restriction should not only apply to the target UOA, but to the average UOA in the set. The average artificial UOA j_a is defined as: $\left(x_{ij_a}, y_{rj_a}\right) = \left(\sum_{j=1}^{n} \frac{x_{ij}}{n}, \sum_{j=1}^{n} \frac{y_{rj}}{n}\right)$, and the following restrictions are added to each linear program in the initial approach (6.2.1):

$$a_r \leq \frac{\left(\sum_{j=1}^{n} \frac{y_{rj}}{n}\right) u_r}{\sum_{r=1}^{s}\left(\sum_{j=1}^{n} \frac{y_{rj}}{n}\right) u_r} \leq b_r, \forall r \quad \text{and} \quad c_i \leq \frac{\left(\sum_{j=1}^{n} \frac{x_{ij}}{n}\right) v_i}{\sum_{i=1}^{m}\left(\sum_{j=1}^{n} \frac{x_{ij}}{n}\right) v_i} \leq d_i, \forall i$$

Thus, when this approach is implemented, each UOA is assessed with four additional constraints for each factor being restricted:

$$
\begin{array}{llcll}
\frac{1}{e_0} = \min \sum_{i=1}^{m} x_{ij_0} v_i & & & & \\
s.t. & & & & \\
 & \sum_{r=1}^{s} y_{rj_0} u_r & = & 1 & \\
\sum_{i=1}^{m} x_{ij} v_i & -\sum_{r=1}^{s} y_{rj} u_r & \geq & 0, & \forall j \\
 & y_{rj_0} u_r & \geq & a_r, & \forall r \\
 & -y_{rj_0} u_r & \geq & -b_r, & \forall r \\
x_{ij_0} v_i - c_i \sum_{i=1}^{m} x_{ij_0} v_i & & \geq & 0, & \forall i \\
-x_{ij_0} v_i + d_i \sum_{i=1}^{m} x_{ij_0} v_i & & \geq & 0, & \forall i \\
 & \left(\sum_{j=1}^{n} \frac{y_{rj}}{n} \right) u_r - a_r \sum_{r=1}^{s} \left(\sum_{j=1}^{n} \frac{y_{rj}}{n} \right) u_r & \geq & 0, & \forall r \\
 & -\left(\sum_{j=1}^{n} \frac{y_{rj}}{n} \right) u_r + b_r \sum_{r=1}^{s} \left(\sum_{j=1}^{n} \frac{y_{rj}}{n} \right) u_r & \geq & 0, & \forall r \\
\left(\sum_{j=1}^{n} \frac{x_{ij}}{n} \right) v_i - c_i \sum_{i=1}^{m} \left(\sum_{j=1}^{n} \frac{x_{ij}}{n} \right) v_i & & \geq & 0, & \forall i \\
-\left(\sum_{j=1}^{n} \frac{x_{ij}}{n} \right) v_i + d_i \sum_{i=1}^{m} \left(\sum_{j=1}^{n} \frac{x_{ij}}{n} \right) v_i & & \geq & 0, & \forall i \\
v_i & & \geq & \varepsilon, & \forall i \\
 & u_r & \geq & \varepsilon, & \forall r
\end{array}
$$

This averaging construct will penalise units with small or large, input or output values, as it imposes a 'majority rule', which is clearly against the spirit of traditional DEA.

The envelopment formulation becomes:

$$\frac{1}{e_0} = \max \theta_0 + \sum_{r=1}^{s} a_r \alpha_r - \sum_{r=1}^{s} b_r \beta_r + \varepsilon \sum_{i=1}^{m} s_i^- + \varepsilon \sum_{r=1}^{s} s_r^+$$

$$s.t.$$

$$\sum_{j=1}^{n} x_{ij}\lambda_j + x_{ij_0}\left(\gamma_i - \sum_{i=1}^{m} c_i\gamma_i - \delta_i + \sum_{i=1}^{m} d_i\delta_i\right) + \left(\sum_{j=1}^{n} \frac{x_{ij}}{n}\right)\left(\gamma_{i_a} - \sum_{i=1}^{m} c_i\gamma_{i_a} - \delta_{i_a} + \sum_{i=1}^{m} d_i\delta_{i_a}\right) + s_i^- = x_{ij_0}, \forall i$$

$$y_{rj_0}\theta_0 - \sum_{j=1}^{n} y_{rj}\lambda_j + y_{rj_0}(\alpha_r - \beta_r) + \left(\sum_{j=1}^{n} \frac{y_{rj}}{n}\right)\left(\alpha_{r_a} - \sum_{r=1}^{s} a_r\alpha_{r_a} - \beta_{r_a} + \sum_{r=1}^{s} b_r\beta_{r_a}\right) + s_r^+ = 0, \forall r$$

$$\theta_0 \quad \text{free}$$

$$\lambda_j \geq 0, \forall j$$

$$\gamma_i, \delta_i, \gamma_{i_a}, \delta_{i_a}, \quad s_i^- \geq 0, \forall i$$

$$\alpha_r, \beta_r, \alpha_{r_a}, \beta_{r_a}, \quad s_r^+ \geq 0, \forall r$$

The obtained targets are then:

$$\sum_{j=1}^{n} x_{ij}\lambda_j^* = x_{ij_0}\left(1 - \gamma_i^* + \sum_{i=1}^{m} c_i\gamma_i^* + \delta_i^* - \sum_{i=1}^{m} d_i\delta_i^*\right) - x_{ij_a}\left(\gamma_{i_a}^* - \sum_{i=1}^{m} c_i\gamma_{i_a}^* - \delta_{i_a}^* + \sum_{i=1}^{m} d_i\delta_{i_a}^*\right) - s_i^{-*}, \forall i$$

$$\sum_{j=1}^{n} y_{rj}\lambda_j^* = y_{rj_0}\left(\theta_0^* + \alpha_r^* - \beta_r^*\right) + y_{rj_a}\left(\alpha_{r_a}^* - \sum_{r=1}^{s} a_r\alpha_{r_a}^* - \beta_{r_a}^* + \sum_{r=1}^{s} b_r\beta_{r_a}^*\right) + s_r^{+*}, \forall r$$

Being a compromise between the first (6.2.1) and second (6.2.2) approaches, the third approach has characteristics of both of them. As in the previous approaches, the efficiency measure, e_0^* will no longer necessarily approximate the inverted radial expansion factor $1/\theta_0^*$. And, the targets are increasingly difficult to interpret: not only are they dependent on the current levels of input and output of the target UOA, they are also dependent on the average value of all the units. As with the previous approach, this alternative can also easily lead to infeasible linear programs, as will be shown later in this chapter.

6.3 A New Classification of Virtual Weights Restrictions

Although inspired by Wong and Beasley's virtual weights restrictions, the virtual weights restrictions used in this thesis are different. It will be shown that the restrictions used present advantages relative to the ones proposed by Wong and Beasley.

Thompson et al. (1990)[6] proposed a classification of absolute weights restrictions into simple absolute weights restrictions and assurance regions. They named ARI the absolute weights restrictions that incorporate into the analysis the relative ordering or values of the inputs or outputs; and ARII restrictions that translate relationships across the input and output divide. In fact, information about a numerical range to translate the importance of the input or output factor as in simple direct restrictions on the multipliers, in general, might be difficult to obtain, and an ordering of preferences might be more suitable.

For this reason, in this section a similar classification for virtual weights restrictions is proposed. That is, to have virtual weights restrictions equivalent to simple absolute weights restrictions, as well as equivalent to assurance regions of type I (ARI) and of type II (ARII). Thus the ordering of preferences by applicants to universities in Chapter 3 are translated by *virtual* weights restrictions of the type I. Indeed, most applicants would not be able to specify explicit weights. The linking of teaching and research outputs produced by a cost group in each university to the inputs available to that cost group, used in Chapter 4, is an example of *virtual* weights restrictions of type II.

All virtual weights restrictions, used in this thesis, can be described by the general set of $w=1...t$ weights restrictions, applying to the target unit:

$$\sum_{i=1}^{m} a_{iw} x_{ij_0} v_i + \sum_{r=1}^{s} b_{rw} y_{rj_0} u_r \geq k_w, \qquad \forall w.$$

This expression encapsulates the three different kinds of virtual weights restrictions, of the new classification presented below.

6.3.1 Simple virtual weights restrictions

Simple virtual weights restrictions involve constraining the virtual of a single factor. This approach is equivalent to using *proportional* virtual weights restrictions applied to output virtuals in an output oriented model. If applied also to input virtuals in an output oriented model, it will only be equivalent for the units that are efficient, and therefore define the frontier. They are of the form:

$$\begin{aligned} a_{iw} x_{ij_0} &\geq k_w, i = i' \\ a_{iw} &= 0, \forall i \neq i' \\ b_{rw} &= 0, \forall r \end{aligned}$$

for restricting the virtual input i'; and

$$\begin{aligned} b_{rw} y_{rj_0} &\geq k_w, r = r' \\ b_{rw} &= 0, \forall r \neq r' \\ a_{iw} &= 0, \forall i \end{aligned}$$

for restricting virtual output r'.

6.3.2 Virtual ARI

Virtual ARI restrictions are equivalent to absolute weights restrictions ARI. They link virtual inputs or outputs to translate an ordering in preference. They are:

$$\sum_{i=1}^{m} a_{iw} x_{ij_0} v_i \geq 0$$
$$b_{rw} = 0, \forall r$$

to link virtual inputs, and

$$\sum_{r=1}^{s} b_{rw} y_{rj_0} u_r \geq 0$$
$$a_{iw} = 0, \forall i$$

to link virtual outputs.

6.3.3 Virtual ARII

Finally, virtual ARII restrictions are equivalent to absolute weights restrictions ARII. They link the input-output divide. They can be translated by:

$$\sum_{i=1}^{m} a_{iw} x_{ij_0} v_i + \sum_{r=1}^{s} b_{rw} y_{rj_0} u_r \geq 0,$$

where at least one $a_{iw} \neq 0$ and one $b_{rw} \neq 0$.

6.4 Virtual Assurance Regions with Factor Linkages

6.4.1 Virtual weights restrictions apply to the target *UOA* j_0

When combinations of different types of virtual weights restrictions are used in a model, the multipliers formulation becomes:

$$
\begin{array}{llll}
\dfrac{1}{e_0} = \min \sum_{i=1}^{m} x_{ij_0} v_i & & & \\
s.t. & & & \\
\sum_{r=1}^{s} y_{rj_0} u_r & = & 1 & \\
\sum_{i=1}^{m} x_{ij} v_i - \sum_{r=1}^{s} y_{rj} u_r & \geq & 0, & \forall j \\
\sum_{i=1}^{m} a_{iw} x_{ij_0} v_i + \sum_{r=1}^{s} b_{rw} y_{rj_0} u_r & \geq & k_w, & \forall w \\
v_i & \geq & \varepsilon, & \forall i \\
u_r & \geq & \varepsilon, & \forall r
\end{array}
$$

The effect on the envelopment formulation is as below:

$$
\begin{array}{llll}
\dfrac{1}{e_0} = \max \theta_0 + \sum_{w=1}^{t} k_w \rho_w + \varepsilon \sum_{i=1}^{m} s_i^- + \varepsilon \sum_{r=1}^{s} s_r^+ & & & \\
s.t. & & & \\
\sum_{j=1}^{n} x_{ij} \lambda_j + \sum_{w=1}^{t} a_{iw} x_{ij_0} \rho_w + s_i^- & = & x_{ij_0}, & \forall i \\
y_{rj_0} \theta_0 - \sum_{j=1}^{n} y_{rj} \lambda_j + \sum_{w=1}^{t} b_{rw} y_{rj_0} \rho_w + s_r^+ & = & 0, & \forall r \\
\theta_0 & & \text{free} & \\
\lambda_j & \geq & 0, & \forall j \\
\rho_w & \geq & 0, & \forall w \\
s_i^- & \geq & 0, & \forall i \\
s_r^+ & \geq & 0, & \forall r
\end{array}
$$

And the targets:

$$\sum_{j=1}^{n} x_{ij}\lambda_j^* = x_{ij_0}\left(1-\sum_{w=1}^{t} a_{iw}\rho_w^*\right) - s_i^{-*}, \forall i$$

$$\sum_{j=1}^{n} y_{rj}\lambda_j^* = y_{rj_0}\left(\theta_0^* + \sum_{w=1}^{t} b_{rw}\rho_w^*\right) + s_r^{+*}, \forall r$$

If the model includes *simple virtual weights restrictions*, where either a minimum and/ or maximum virtual is imposed for some or all factors, then the efficiency score e_0^* will no longer necessarily approximate the inverted radial expansion factor $1/\theta_0^*$, as the objective function contains a term with the new variable ρ_w to be maximised. On the other hand, if the model contains only *virtual weights restrictions of the type ARI and ARII*, the efficiency score will converge to be $1/\theta_0^*$. As for the targets, their interpretation is easier than in the models with *proportional* virtual weights restrictions. Either an improvement or deterioration of current levels of the factors is possible, but in any case they can still be interpreted as a contraction or expansion of the current levels of the factors of the target unit.

The same problems discussed, when *proportional* virtual weights restrictions apply only to the target unit, occur. Restrictions applying to all units can be envisaged, as in the next section.

6.4.2 *Virtual weights restrictions apply to* all *UOA* j

The multipliers formulation with the virtual weights restrictions applying to all UOAs are as below:

$$
\begin{aligned}
&\frac{1}{e_0} = \min \sum_{i=1}^{m} x_{ij_0} v_i \\
&s.t. \\
&\qquad\qquad \sum_{r=1}^{s} y_{rj_0} u_r \quad = \quad 1 \\
&\sum_{i=1}^{m} x_{ij} v_i \quad - \sum_{r=1}^{s} y_{rj} u_r \quad \geq \quad 0, \quad \forall j \\
&\sum_{i=1}^{m} a_{iw} x_{ij} v_i \quad + \sum_{r=1}^{s} b_{rw} y_{rj} u_r \quad \geq \quad k_w, \quad \forall w, j \\
&v_i \quad \geq \quad \varepsilon, \quad \forall i \\
&u_r \quad \geq \quad \varepsilon, \quad \forall r
\end{aligned}
$$

The effect on the envelopment formulation would be as below:

$$
\begin{aligned}
&\frac{1}{e_0} = \max \theta_0 + \sum_{j=1}^{n} \sum_{w=1}^{t} k_w \rho_{wj} + \varepsilon \sum_{i=1}^{m} s_i^- + \varepsilon \sum_{r=1}^{s} s_r^+ \\
&s.t. \\
&\sum_{j=1}^{n} x_{ij} \lambda_j \quad + \sum_{j=1}^{n} \sum_{w=1}^{t} a_{iw} x_{ij} \rho_{wj} \quad + s_i^- \quad = \quad x_{ij_0}, \quad \forall i \\
&y_{rj_0} \theta_0 \quad - \sum_{j=1}^{n} y_{rj} \lambda_j \quad + \sum_{j=1}^{n} \sum_{w=1}^{t} b_{rw} y_{rj} \rho_{wj} \quad + s_r^+ \quad = \quad 0, \quad \forall r \\
&\theta_0 \quad \text{free} \\
&\lambda_j \quad \geq \quad 0, \quad \forall j \\
&\rho_{wj} \quad \geq \quad 0, \quad \forall w, j \\
&s_i^- \quad \geq \quad 0, \quad \forall i \\
&s_r^+ \quad \geq \quad 0, \quad \forall r
\end{aligned}
$$

And the targets:

$$\sum_{j=1}^{n} x_{ij}\lambda_j^* = x_{ij_0} - \sum_{j=1}^{n}\sum_{w=1}^{t} a_{iw} x_{ij} \rho_{wj}^* - s_i^{-*}, \forall i$$

$$\sum_{j=1}^{n} y_{rj}\lambda_j^* = y_{rj_0}\theta_0^* + \sum_{j=1}^{n}\sum_{w=1}^{t} b_{rw} y_{rj} \rho_{wj}^* + s_r^{+*}, \forall r$$

As in the previous section, if the model includes *simple virtual weights restrictions*, where either a minimum and/ or maximum virtual is imposed for some or all factors, then the efficiency score e_0^* will no longer necessarily approximate the inverted radial expansion factor $1/\theta_0^*$. On the other hand, if the model contains only *virtual weights restrictions of the type ARI and ARII*, the efficiency score will converge to be $1/\theta_0^*$.

However, the interpretation of the targets as a contraction or expansion of the current levels of inputs or outputs, depending on the controller's preferences translated by the weights restrictions imposed, no longer applies. The expression of the targets for the UOA under analysis has a new term, that not only depends on the current levels of the factor for the target unit but also for all the other units. However, it is still an improvement from the targets obtained from the use of *proportional* virtual weights restrictions applying to all units, in that the target for the factor under analysis does not depend on the virtual limits imposed on all the other factors.

In conclusion, the use of *virtual assurance regions* applying to the target UOA j_0 seems to be the best approach. This approach was widely used in this thesis. It allows for the natural representation of preference structures; linkages between inputs and outputs translating established patterns; and finally, the meaning of the efficiency score and targets are most easily interpreted.

6.5 Using Absolute Weights Restriction with a 'Virtual' Meaning

A problem with virtual weights restrictions is that they are UOA specific. Allen et al (1997)[7] have suggested that virtual weights restriction, as proposed by Wong and Beasley's second alternative (6.2.2), could be converted into absolute weights restrictions, in the following manner:

When considering a lower bound on output r, of a_r, such as $a_r \leq y_{rj}u_r, \forall j$, clearly the virtual restriction corresponding to the UOA with the lowest output level can be binding. Similarly, if a virtual upper bound restriction were imposed on an output r, of b_r, such as $y_{rj}u_r \leq b_r, \forall j$, the binding virtual restriction would be the one corresponding to the UOA with the largest input or output level. The authors conclude that a more economical approach to Wong and Beasley's would be to add only the required binding absolute restrictions. However, this alternative also presents some problems.

Consider the example of a factor, whose importance should, according to the controller, be between 5% and 15%. If the maximum value for this factor is 100, and the minimum 1 (as is the case in some of *The Times* league table factors, for example), the following expression is derived:

$$5\% \leq y_{rj}u_r \leq 15\%, \forall j$$

Using Allen et al's[7] approach, it would lead to the following:

$$\frac{0.05}{\min(y_{rj})} \leq u_r \leq \frac{0.15}{\max(y_{rj})} \Leftrightarrow 0.05 \leq u_r \leq 0.0015\,!$$

In conclusion, the alternative will easily lead to infeasible results for some 'intuitive' set of bound on the virtuals. In fact, for the use of *proportional* virtual weights restrictions to be feasible for all units in each linear program, the bounds have to be carefully chosen, such that:

$$\frac{a_r}{\min(y_{rj})} \le u_r \le \frac{b_r}{\max(y_{rj})}$$

is feasible for all the units. This will mean:

- setting up the lower limit a_r, and then calculate the upper limit b_r as follows:

$$a_r \le y_{rj} u_r \le \frac{\max(y_{rj})}{\min(y_{rj})} \cdot a_r \le b_r, \forall r, j$$

- or, setting up the upper limit b_r, and then a_r is calculated as follows:

$$a_r \le \frac{\min(y_{rj})}{\max(y_{rj})} \cdot b_r \le y_{rj} u_r \le b_r, \forall r, j.$$

Going back to our previous example, if a_r were set at the 5% level, then b_r would have to be at least 500%, for the restriction to be feasible for all UOAs in section 6.2.2. If b_r were set at the 15% level, then a_r could not exceed 0.15%. These results are clearly no longer intuitive, as promoted by Wong and Beasley[4].

A similar reasoning for approach 6.2.3, where the restrictions apply to both target and average unit simultaneously, can be made:

$$\begin{cases} \dfrac{a_r}{y_{rj_0}} \le u_r \le \dfrac{b_r}{y_{rj_0}} \\ \dfrac{a_r}{y_{rj_a}} \le u_r \le \dfrac{b_r}{y_{rj_a}} \end{cases} \Leftrightarrow \begin{cases} y_{rj_0} \le y_{rj_a} \\ \dfrac{a_r}{y_{rj_0}} \le u_r \le \dfrac{b_r}{y_{rj_a}} \end{cases} \vee \begin{cases} y_{rj_0} \ge y_{rj_a} \\ \dfrac{a_r}{y_{rj_a}} \le u_r \le \dfrac{b_r}{y_{rj_0}} \end{cases}$$

for each linear program.

Two situations arise, either the level of the output for the target unit is less than the average, or greater than the average. However, for all linear programs to be feasible, in either case, the same conclusion as for 6.2.2 is reached:

$$a_r \le \frac{\min(y_{rj})}{\max(y_{rj})} \cdot b_r \, .$$

Wong and Beasley[4] were optimistic that obtaining the limits, a_r and b_r corresponding to value judgements, was 'not too difficult', but the multicriteria decision making (MCDM) perspective on the use of DEA disagrees (see, for instance, Belton and Vickers, 1993[8]; Stewart, 1996[9]; Belton and Stewart, 1997[10]). Recently, Belton and Stewart[10] on a discussion on the interactions between MCDM and DEA, point out that the DEA field could learn from the extensive experience of MCDM in eliciting and working with value judgements. They consider that it is difficult to set meaningful bounds on the weights, specially in the case of multiple input - multiple output problems, except in terms of ratios. The realisation that using explicit boundaries for weights is a difficult task, has favoured the use in this thesis of virtual ARs, rather than the simple direct restrictions on the virtuals. Moreover 'intuitive' limits expressed as a percentage, as seen, can easily lead to infeasibility in approaches 6.2.2 and 6.2.3 above.

6.6 DEA and MCDM

The primary focus of DEA is that of comparing a number of UOAs which differ in the quantities of inputs which they consume and in the outputs they produce. The aim is not, in general, to select a UOA, although it can be. The intention is, normally, to identify those UOAs which are inefficient, and to assess where the inefficiencies arise. Nevertheless, the efficiency score also implies some form of rank ordering of the UOAs. This characteristic was used in Thompson et al[11] to solve a choice problem using DEA.

By contrast in MCDM, it is assumed that there are a number of *alternatives* between which the controller has to decide, where each alternative is described by its performance on each of a number of *criteria*. The aim of MCDM is to provide support to the controller in the process of making the choice between alternatives, which may, as in DEA, lead to some form of preference ranking. Several authors have shown that the MCDM and DEA formulations coincide if inputs and outputs are viewed as criteria for evaluating UOAs, with minimisation of inputs and/ or maximisation of outputs as associated objectives (see Doyle and Green, 1991[11]; Belton and Vickers, 1993[8]; and Stewart, 1996[9]). The differences arise on how to suitably constrain the weight flexibility in DEA, if the controller's judgemental values are to be meaningfully incorporated.

Stewart (1996)[9] has demonstrated the equivalence of the conventional DEA ratio form of efficiency definition, and a distance to the Pareto frontier in a linear value function model, which lends itself better to the value elicitation and sensitivity studies which have been developed within the MCDM context. In MCDM terms, the aggregate output and aggregate input are two high level criteria, of which the individual inputs and outputs are sub-criteria. From either DEA or MCDM, Pareto optimality is the important consideration. Hence, if for any UOA j (or 'alternative') there exists another UOA j_0 such that $x_{ij} \geq x_{ij_0}, \forall i$ and $y_{rj} \leq y_{rj_0}, \forall r$, then j is dominated by j_0; that is j_0 is more efficient (thus 'preferred to') j. It is useful, therefore, to examine for each UOA j_0 whether it is possible to find positive weights such that:

$$\sum_{r=1}^{s} y_{rj_0} u_r - \sum_{i=1}^{m} x_{ij_0} v_i \geq \sum_{r=1}^{s} y_{rj} u_r - \sum_{i=1}^{m} x_{ij} v_i, \forall j$$

in which case UOA j_0 is efficient. The answer to this question can be given by the following linear program:

$$\begin{aligned}
&\min \sum_{j=1}^{n} D_j \\
&s.t. \\
&\sum_{r=1}^{s} \left(y_{rj_0} - y_{rj}\right) u_r - \sum_{i=1}^{m} \left(x_{ij_0} - x_{ij}\right) v_i + D_j \geq 0, \forall j \\
&\sum_{r=1}^{s} u_r + \sum_{i=1}^{m} v_i = 1 \\
&u_r, v_i, D_j \geq 0
\end{aligned}$$

If the optimal value of the objective function is zero, then there exists a linear value function for which the target UOA is optimal. The above model classifies UOAs as efficient or not, but does not include any definition of the degree of efficiency. Though, if the aim of the analysis is ultimately to seek a single 'best' UOA (a choice problem, as in Chapter 3, *The Applicant's Perpective*), then the degree of efficiency is irrelevant, since only efficient UOAs can be optimal.

Other differences from the traditional DEA approach are:

Firstly, the MCDM approach proposes that the input and output measures should be scaled, in such a way that the maximum value is some specified value, for example 100 units. Such scaling would facilitate the interpretation of the results, as it avoids the problem with directly restricting the multipliers, i.e. *absolute* weights restrictions, in that they are no longer dependent on the units of measurement of the inputs and outputs. Secondly, they also propose the standardisation of the weights ($\sum_{r=1}^{s} u_r + \sum_{i=1}^{m} v_i = 1$, for instance) in order to facilitate judgement. With standardised scales, it is easier for the controller to consider possible trade-offs between the factors. These two measures, in fact, pre-empt the need for the use of virtual weights restrictions.

Traditional DEA is a technique for extracting the maximum *objective* information concerning differences in performance of the different units par excellence. MCDM, instead, has developed a considerable body of knowledge concerning the assessment of *value judgements* and the role of sensitivity analysis in supporting such judgements.

The inclusion of weights restrictions is an exercise in *value judgement*, but it still only generates one single efficiency measure. This is in contradiction with the uncertainty in the mind of the controller as to what weights best reflect the relative importance of the individual inputs and outputs, and a single measure cannot capture the extent of this uncertainty.

Doyle and Green[11] proposed the use of cross-efficiency in DEA, to overcome the problem. Cross-efficiencies are the efficiencies determined for each UOA by using the optimal weights from the other units in turn. The average of any given UOA can be thought of as the average evaluation from each of the other UOAs. Cross-efficiencies can be used also in a restricted model, and will give a measure of the 'stability' of the results. Thus a unit which has high cross-efficiencies will be a good all round unit, whereas a unit with low cross-efficiencies, albeit efficient, might be performing very well in a particular factor, but not so good in other factors.

Belton and Vickers[8] suggested an ad-hoc sensitivity analysis based on a visual interactive approach, involving the controller. The controller can change graphically the weights restrictions applying to the different factors in the problem, and then visualise how that affects the positioning of a unit in relation to the frontier of an XY scatter plot (aggregate input by the aggregate output) of all UOAs under analysis.

Stewart[9] proposed a more systematic approach, that also overcomes the acknowledged problem in DEA of multiple solutions corresponding to alternative weights vectors. Again the approach will suit both unrestricted, and restricted DEA models. It is based on the fact that any probability distribution over weights implies distribution for the efficiencies of each UOA. These distributions are easily obtained by Monte Carlo methods. The approach allows users to obtain a direct picture of what the uncertainties in weights imply. The upper bound of the distribution of efficiencies still give the standard DEA solution, while the remainder of the distribution gives a clearer picture of how extreme or implausible the weights need to be for any given degree of efficiency to be attributed to a particular UOA. For example, two UOAs might both be efficient, but one might be efficient for almost all allowable weights, whereas the other is efficient only for weights in a subset of very low probability. An even more interesting insight might arise if one UOA is efficient, but only for unlikely weight combinations, while another is inefficient, but with an efficiency never far below 1. This awareness would be difficult to obtain from the standard DEA analysis.

6.7 Concluding Remarks

In this chapter a new classification for virtual weights restrictions is proposed. In general, virtual weights restrictions are easier to interpret than *proportional* virtual weights restrictions. Because of the problems described, in this thesis the concept of assurance regions, rather than imposing bounds on virtuals is often preferred. It is shown how the concept of assurance regions in absolute weights restrictions can be imported into virtual weights restrictions. It is concluded that the use of *virtual assurance regions* applying to the target UOA j_0 is the best approach for the natural representation of preference structures and to translate established patterns across the input-output divide. Also, the meaning of the efficiency score and targets in this approach are most easily interpreted. For this reason, this was the approach used throughout this thesis.

Soon after the first applications of the DEA technique to real situations, people were faced with the problem that the efficiency of certain UOAs was dependent upon rather unlikely sets of weights. Restrictions on the weights were thus introduced into 'classical' DEA to curb its total flexibility, and obtain more sensible results. A problem with placing absolute restrictions on the weights though, is that these will be dependent on the units of measurement of the inputs and outputs, and will consequently be difficult to devise.

The latter problem was overcome with the introduction of virtual weights restrictions. Virtuals are dimensionless and give a measure of the importance of the factor in the measurement of efficiency. It is easier to think of meaningful virtuals, rather than absolute weights. This characteristic becomes even more important when there is an external decision maker to the UOA, the controller, who is trying to impose a preference structure on the evaluation of the UOAs.

The consequences of the use of absolute weights restrictions on the efficiency measurement, targets and peer units has been extensively reviewed by Allen et al.[5], but the consequences of the use of virtual weights restrictions has been less explored in the literature.

Wong and Beasley[4] proposed a judicious method of constraining the virtuals based on proportions: the controller can impose a range on the virtual factors in terms of percentage 'importance'. The approach is not without fault, however. Virtual weights restrictions are UOA specific, and one has to decide to which units the restrictions apply. Applying them solely on the target UOA, means that the other UOAs are being imposed meaningless virtuals. At first thought, imposing weights restrictions on all the units might solve the problem. However, in this chapter it is shown that the imposition of 'unfair' restrictions also happens in this case. Moreover, this approach is computationally expensive, and more worrying, it very often leads to infeasibility. A third compromise approach does not seem to alleviate the latter problem. Moreover, the targets obtained for inefficient units are increasingly difficult to interpret.

The use of virtual weights restrictions applying to the target unit only, as opposed to *proportional* virtual weights restrictions, are advocated in this thesis. They still present the problem that the target unit is evaluated in a more lenient manner than the other units. However, the interpretation of the targets is more meaningful and it does not lead to the infeasibility problems of the use of *proportional* weights restrictions applying to all units.

Allen et al's[7] suggestion of transforming proportional virtual weights restrictions into absolute ones often leads to infeasibility. However, the idea would be very useful in transforming simple virtual weights restrictions, when only lower or upper bounds have been determined for each factor, but not both simultaneously.

The perspective of traditionally MCDM authors on DEA has produced some recent insights into the problem of weights restrictions. DEA used as an MCDM tool is able to screen a number of alternatives in an MCDM problem, by extracting a maximum of 'objective' information. On the other hand, the MCDM researchers suggest that the introduction of bounds on weights in DEA could benefit by the body of understanding concerning the assessment of value judgements and the role of sensitivity analysis in supporting such judgements, developed in the MCDM field.

6.8 References

[1] Thompson RG, Singleton FD, Thrall RM and Smith BA (1986). Comparative site evaluations for locating a high-energy physics lab in Texas. *Interfaces* **16**: 35-49.

[2] Dyson RG and Thanassoulis E (1988). Reducing weight flexibility in data envelopment analysis. *J Opl Res Soc* **39**: 563-576.

[3] Charnes A, Cooper WW, Huang ZM and Sun DB (1990). Polyhedral cone-ratio DEA models with an illustrative application to large commercial banks. *J Econometrics* **46**: 73-91.

[4] Wong Y-HB and Beasley JE (1990). Restricting weight flexibility in data envelopment analysis. *J Opl Res Soc* **41**: 829-835.

[5] Allen R, Athanassopoulos AD, Dyson RG and Thanassoulis E (1997). Weights restrictions and value judgements in data envelopment analysis: evolution, development and future directions. *Ann Ops Res* **73**: 13-34.

[6] Thompson RG, Langemeier LN, Lee C-T, Lee E, Thrall RM (1990). The role of multiplier bounds in efficiency analysis with application to Kansas farming. *J Econometrics* **46**: 93-108.

[7] Allen R, Dyson RG, Santos AR and Shale E (1997). Pitfalls and protocols in the application of DEA. Presented at the *39th Annual Conference of the Operational Research Society*. September 1997, Bath, UK.

[8] Belton V and Vickers SP (1993). Demystifying DEA - a visual interactive approach based on multiple criteria analysis. *J Opl Res Soc* **44**: 883-896.

[9] Stewart TJ (1996). Relationships between data envelopment analysis and multicriteria decision analysis. *J Opl Res Soc* **47**: 654-665.

[10] Belton V and Stewart TJ (1997). DEA and MCDM: competing or complementary approaches? Research Paper, University of Strathclyde.

[11] Doyle JR and Green RH (1991). Comparing products using data envelopment analysis. *Omega* **19**: 631-638.

7. SUMMARY AND CONCLUSIONS

7.1 Summary

The objective of Chapter 1 was to give the background of the higher education sector in the UK at present and its history. It introduces the drive for performance measurement in higher education, and the motivation for the dissertation: bringing in the stakeholders perspectives on performance measurement in UK universities using data envelopment analysis. The method data envelopment analysis is then described.

In Chapter 2 the traditional use of performance indicators and peer assessment is reviewed and the use of DEA, instead of parametric techniques, is justified. The review of the literature on the use of DEA in performance measurement in universities catalogues three kinds of studies: comparing universities; comparing university departments; and appraisals of individuals within the university (staff and students). The opportunity to use DEA in a somewhat different way than previously is identified. The novel proposed framework integrates in the same analysis the perspectives of three different levels of stakeholders. Firstly, the perspective of the applicant in the process of choosing a university to apply to; secondly, the perspective of the State that funds and evaluates university performance; and finally the institutional perspective, which includes the views of the latter two external evaluations placed on it.

In Chapter 3 the applicant's perspective is presented. The use of DEA in university selection is compared to existing methods and a new approach is proposed that uses DEA. The approach devised recognises the different values of students and is empirically tested in a case study at a comprehensive school. In a first part, the process of decision making is looked at in a retrospective way, making use of students that had already undergone the experience of applying to university. In a second part, the insight gained from the first part is used to build a model that would help students in the process of choosing a university to apply to. Chapter 3 clearly deals with a choice problem, and the link with MCDM is first approached. Finally, a comprehensive decision support system that includes DEA for university selection is arrived at.

Chapter 4 also deals with an external summative evaluation on university performance, from the perspective of the State. The relationship between the State and higher education over time is described, the current operational model explained and the future trends outlined. In order to measure performance, according to the mission and objectives of the state/ funding councils, a review of their three main remits is undertaken. The contribution of DEA to inform the State/ funding councils in their remit is then discussed. Two different kinds of models are proposed: to inform the assessment of quality, and to inform the resource allocation separately for teaching and research, as is the current practice of the funding councils. The conceptual link between the two different kinds of models is then described. An assessment of the total performance of the university then follows, i.e. with teaching and research together, from two different perspectives: accountability and autonomy from State funds; and a more traditional evaluation of efficiency. The single difference between the two approaches is the positioning of earned income in the input/ output divide. These two approaches reflect two *different performances*, which have different ideological supporters. Finally, a technical problem of taking account of subject mix factor in the measurement of performance is dealt with, by linking the input/ output divide by means of weights restrictions. An application to the university sector for the academic year 1995/96 is incomplete due to problems of data availability, which have only allowed for some of the models proposed to be solved.

Chapter 5 shows how institutions can turn performance measurement to their own benefit, by using it as a formative exercise to understand the different expectations of them, by the two previous external evaluations. A methodology for institutional performance management is proposed that takes into account the external/ internal interfaces: the applicant/ institution, and state/ institution interfaces. The methodology is illustrated with an application to the University of Warwick.

Virtual weights restrictions are widely used in this thesis, a reflection on its uses is offered in Chapter 6. The reasons for mainly using virtual weights restrictions instead of absolute weights restrictions are explained. However, although the consequences of the use of absolute weights restrictions on the efficiency score and targets of a DEA model have been explored elsewhere, the same was not true for the use of restrictions on the virtuals. A new classification of virtuals weights restrictions is presented. The use of *proportional* weights restrictions is reviewed, and the reasons for using simple virtual weights and virtual assurance regions in this thesis are justified. Alternatives to using virtual weights restrictions are considered, namely using absolute weights restrictions with a virtual meaning. In recognition of levels of decision making at the unit, and external to the unit, the use of the terms unit of assessment (UOA) and controller is proposed. The relationship between DEA and MCDM in this domain is elaborated upon, following an initial reference in Chapter 3.

In the next sections the conclusions derived from the thesis and a discussion of the methodology used in this thesis are presented, as well as suggestions for further research.

7.2 Conclusions

7.2.1 The importance of bringing in the perspectives of different stakeholders

The first main conclusion of the present thesis was that previous DEA studies in the area of performance measurement in universities failed to address the existence of different stakeholders interested in performance measurement in universities, and the emergence of interactions between the different perspectives. The approach proposed sought to model the different perspectives and their interactions using DEA with weights restrictions. The stakeholders multi-perspective approach is a contribution to the methodology of organisational performance measurement.

7.2.2 The contribution of DEA in choice problems

DEA with weights restrictions is an appropriate methodology to deal with choice problems. Chapter 3 shows the contribution of DEA in choice problems, by focusing on the evaluation of universities from the perspective of the prospective student. DEA with weights restrictions is capable of recognising differences in institutional mission through allowing variability in the weights, and at the same time, it recognises the priorities of applicants, by the introduction of weights restrictions. DEA is also more flexible, as opposed to popular methodologies, such as The Times', that assign common values to all decision makers. Applicant's to universities differ in their objectives and emphases on different criteria when evaluating universities. DEA allows for customisation to suit individual decision makers. DEA also proved to be an effective technique for the aggregation of a considerable amount of information, as part of a more comprehensive decision support system.

7.2.3 Subject mix handled by virtual ARII

Handling subject mix in the context of university performance measurement has long been a problem. A procedure is required which ensures that universities with low cost subjects do not have an unfair advantage in the assessment of performance. In the past, DMUs have been divided into different comparable sets, and assessed separately. In Chapter 4 a solution for taking into account subject mix in universities is arrived at by linking inputs and outputs across the input/ output divide. This represents a novel application of the method devised by Thanassoulis et al.[1], which prevents units from taking undue advantage of weight flexibility contrary to the known links between certain inputs and outputs. It is the first time it is used in the context of university performance measurement. Also, the concept was translated into virtual weights restrictions, instead of the absolute weights restrictions used in the original application.

7.2.4 Inadequate data availability policy

In Chapter 4 it is demonstrated that DEA could inform peer judgement in the assessment of quality, bringing in a much desired objectivity into the current process. Moreover DEA could inform the allocation of resources on a value-for-money approach. A theoretical link between DEA models of quality assessment and resource allocation, which would be an improvement on the current methodologies used by the funding councils, is proposed.

The proposed methodology could not be tested because the performance measures necessary were not available. The results of the models that could be run, relating to the total performance of universities (i.e. teaching and research) are not completely robust, because different measures are taken at different levels of aggregation. In Chapter 5, when considering the institutional perspective, only analysis at subject level is meaningful, however most relevant data is not to be found at subject level.

The policy of data availability at the time of the thesis fails in three distinct points:

- the appropriateness of the measures available;
- the aggregation at subject level; and
- the aggregation at cost level.

As to the appropriateness of the measures available, it is for the universities to agree on a set of relevant measures that will help them monitor their performance. So far, the university sector has failed to reach a consensus on the data that should be collected and made available. However, if universities do not take ownership of the issue of performance measures, external evaluators will. In that case, it will be difficult for institutions to argue their case. As to the level of aggregation regarding cost levels, the funding councils have already planned to use a consistent aggregation at cost code level for all the data they collect. In the future, it will thus be possible to make the analysis proposed in this thesis at cost level and subject level.

7.2.5 A more appropriate way of comparing departments within a university

This thesis presents a novel approach for the comparison of departments of different subjects within the same institution. Departing from previous studies, which compare different subject departments directly, in this approach departments are firstly judged against those of the same discipline in other universities to ensure comparability; and then their performance, thus assessed, compared to the other departments of the same university. The methodology proposed for institutional performance management, enables the institution to allocate resources among departments taking into consideration, not only its own mission and objectives, but also how those interact with the two external evaluations placed on it. It basically integrates the three different levels of stakeholders perspectives.

7.2.6 The usefulness of virtual assurance regions

Following a critique of Wong and Beasley's proposed method for constraining the virtuals in DEA, a new classification scheme for virtual weights restrictions is presented, which brings the concept of assurance regions into virtual weights restrictions. It is shown that the use of simple virtual restrictions and virtual assurance regions as used in this thesis are preferable to the use of the more generally advocated proportional virtual weights restrictions.

It is concluded that the use of *virtual assurance regions* applying to the target UOA j_0 is the best approach for the natural representation of preference structures and to translate established patterns between the input-output divide. Also, the meaning of the efficiency score and targets in this approach are more easily interpreted.

7.3 Discussion of Methodology

This thesis is eclectic in its nature. It not only aims to contribute to the broad area of organisational performance measurement, but also to the more technical research area of DEA. The hybrid essence of the thesis, however, does not prevent its coherence. The running theme of the dissertation is that the application of DEA to concrete situations, necessarily confronts the user of DEA with the organisational implications of a performance measurement exercise. Both aspects, technical and organisational, should be dealt with in any assessment of performance. Thus, this thesis is concerned with establishing a theoretical framework for the measurement of performance in universities using DEA.

Applied research of this nature raises some issues. What distinguishes this sort of research from consultancy? What are the ethical implications of taking the viewpoint of the different stakeholders involved? Is the researcher working for any of the stakeholders?

The applied research in this thesis is indeed different from consultancy work. None of the stakeholders mentioned in the thesis has commissioned any of the work, nor do they necessarily condone its results. The most fundamental difference, however, is the generalisability of the findings. The stakeholders multi-perspective approach is a contribution to the methodology of organisational performance measurement, as mentioned above. The approach can be used in other higher education systems, with other stakeholders. Moreover, it can be used in other multi-level organisations undergoing performance measurement exercises.

7.4 Suggestions for Further Research

The differences in performance of different universities, depending on age, location, and structure of funding should be assessed when better data becomes available. The results would have possibly important consequences for state policy decisions. It would also be interesting to examine performance over time, bringing in novel research in dynamic DEA.

As to the incorporation of quality into DEA models, more research is needed on how to deal with index measures, which are typically used to describe qualitative aspects of performance. If some qualitative aspects cannot be meaningfully dealt within the DEA model itself, the relationship between quantitative and qualitative analyses and how to integrate the two in a coherent model should be further explored .

Finally the relationship between DEA and MCDM holds much promise, *specially in* eliciting weights from the decision maker when judging what we would call units of assessment, rather than DMUs, in choice problems.

Given the multiplicity of applications in the 20 years after the original DEA publication by Charnes, Cooper and Rhodes in 1978, more is to be expected in DEA research, as applications to real problems demand even more answers, and thus push the boundaries of theoretical knowledge.

7.5 References

[1] Thanassoulis E, Boussofiane A and Dyson RG (1995). Exploring output quality targets in the provision of perinatal care in England using data envelopment analysis. *Euro J Opl Res* **80**: 588-607.